COMPUTERS AND THEIR ROLE IN ENERGY RESEARCH: CURRENT STATUS AND FUTURE NEEDS

HEARINGS

BEFORE THE

SUBCOMMITTEE ON
ENERGY DEVELOPMENT AND APPLICATIONS

AND THE

SUBCOMMITTEE ON
ENERGY RESEARCH AND PRODUCTION

OF THE

COMMITTEE ON
SCIENCE AND TECHNOLOGY
U.S. HOUSE OF REPRESENTATIVES

NINETY-EIGHTH CONGRESS

FIRST SESSION

JUNE 14 AND 15, 1983

[No. 20]

Printed for the use of the
Committee on Science and Technology

U.S. GOVERNMENT PRINTING OFFICE
23-462 O WASHINGTON : 1983

H701-74

COMMITTEE ON SCIENCE AND TECHNOLOGY

DON FUQUA, Florida, *Chairman*

ROBERT A. ROE, New Jersey
GEORGE E. BROWN, JR., California
JAMES H. SCHEUER, New York
RICHARD L. OTTINGER, New York
TOM HARKIN, Iowa
MARILYN LLOYD, Tennessee
DOUG WALGREN, Pennsylvania
DAN GLICKMAN, Kansas
ALBERT GORE, JR., Tennessee
ROBERT A. YOUNG, Missouri
HAROLD L. VOLKMER, Missouri
BILL NELSON, Florida
STAN LUNDINE, New York
RALPH M. HALL, Texas
DAVE McCURDY, Oklahoma
MERVYN M. DYMALLY, California
PAUL SIMON, Illinois
NORMAN Y. MINETA, California
RICHARD J. DURBIN, Illinois
MICHAEL A. ANDREWS, Texas
BUDDY MacKAY, Florida
TIM VALENTINE, North Carolina
HARRY M. REID, Nevada
ROBERT G. TORRICELLI, New Jersey
FREDERICK C. BOUCHER, Virginia

LARRY WINN, JR., Kansas
MANUEL LUJAN, JR., New Mexico
ROBERT S. WALKER, Pennsylvania
WILLIAM CARNEY, New York
F. JAMES SENSENBRENNER, JR., Wisconsin
JUDD GREGG, New Hampshire
RAYMOND J. McGRATH, New York
JOE SKEEN, New Mexico
CLAUDINE SCHNEIDER, Rhode Island
BILL LOWERY, California
ROD CHANDLER, Washington
HERBERT H. BATEMAN, Virginia
SHERWOOD L. BOEHLERT, New York
ALFRED A. McCANDLESS, California
TOM LEWIS, Florida

J. H. Poore, *Executive Director*
Robert C. Ketcham, *General Counsel*
Regina A. Davis, *Administrator*
David S. Jeffrey, *Minority Staff Director*

Subcommittee on Energy Development and Applications

DON FUQUA, Florida, *Chairman*

RICHARD L. OTTINGER, New York
TOM HARKIN, Iowa
RALPH M. HALL, Texas
PAUL SIMON, Illinois
RICHARD J. DURBIN, Illinois
FREDERICK C. BOUCHER, Virginia
ROBERT A. ROE, New Jersey
JAMES H. SCHEUER, New York
MARILYN LLOYD, Tennessee
ROBERT A. YOUNG, Missouri
BILL NELSON, Florida
DAVE McCURDY, Oklahoma
NORMAN Y. MINETA, California

F. JAMES SENSENBRENNER, JR., Wisconsin
CLAUDINE SCHNEIDER, Rhode Island
BILL LOWERY, California
HERBERT H. BATEMAN, Virginia
TOM LEWIS, Florida
WILLIAM CARNEY, New York
JUDD GREGG, New Hampshire

Subcommittee on Energy Research and Production

MARILYN LLOYD, Tennessee, *Chairman*

ROBERT A. ROE, New Jersey
ROBERT A. YOUNG, Missouri
STAN LUNDINE, New York
RICHARD L. OTTINGER, New York
RALPH M. HALL, Texas
TIM VALENTINE, North Carolina

ROBERT S. WALKER, Pennsylvania
ROD CHANDLER, Washington
MANUEL LUJAN, JR., New Mexico
BILL LOWERY, California

CONTENTS

WITNESSES

June 14, 1983:

Page

- Larry W. Sumney, executive director, Semiconductor Research Corp., James Hill, president and chief executive officer, Denelcor, Inc.; and John F. Carlson, vice president, Government relations, Cray Research, Inc. 4
- Robert S. Cooper, Director, Defense Advanced Research Projects Agency; Alvin W. Trivelpiece, Director, Office of Energy Research, Department of Energy; Jack L. Kerrebrock, Associate Administrator, Office of Aeronautics and Space Technology, National Aeronautics and Space Administration 57

June 15, 1983:

- John Roberts, Argonne National Laboratory, Argonne, Ill.; William A. Lester, Jr., principal investigator, Lawrence Berkeley Laboratory, Berkeley Calif., and A. Carl Haussmann, Associate Director, Lawrence Livermore National Laboratory, Livermore, Calif. 98
- Sidney Fernbach, consultant; George Olson, Colorado State University Research Foundation, Fort Collins, Colo., and Robert McCrory, University of Rochester, Laboratory for Laser Energetics, Rochester, N.Y 325

Appendix: Additional material submitted for the record 378

COMPUTERS AND THEIR ROLE IN ENERGY RESEARCH—CURRENT STATUS AND FUTURE NEEDS

TUESDAY, JUNE 14, 1983

House of Representatives, Committee on Science and Technology, Subcommittee on Energy Development and Applications, and Subcommittee on Energy Research and Production,

Washington, D.C.

The joint session met, pursuant to call, at 2 p.m., in room 2318, Rayburn House Office Building, Hon. Don Fuqua (chairman of the Subcommittee on Energy Development and Applications) presiding.

Mr. Fuqua. The joint meeting will come to order.

In the last 30 years, the scientific and engineering accomplishments of our society have been paced by the computational power that has been available. In the early decades of this computer age, the speed and memory of large numerical computers increased by factors of 1,000 every 10 years. The rapid development of supercomputers was the result of productive cooperation between government programs and the technically based entrepreneurs of the computer industry.

This country still has the corner on both the supercomputer market and on the underlying technical base, but the pace of development and application has slowed in recent years. And it is this slowed rate of growth which concerns us in these hearings today and tomorrow. It is a concern because it may allow others to preempt our lead in an area with special impact on our national security and international trade position.

But more important to the Science and Technology Committee are the effects, or lack of them, on the rate of scientific progress. These hearings will focus primarily on the Department of Energy's role in that progress. The department's research and development programs have played a special role in nurturing the development and utilization of supercomputers, and the future of its programs will depend on continued improvements.

It is clear that it is time to revitalize the partnership between government and industry to hasten the development of the next generation of supercomputers.

Today we will hear from industry and government policymakers. Tomorrow we will hear from members of national laboratories and the university community, whose research depends on access to large scale machines. We look forward to your testimony.

At this time, I would like to ask if the chairman of the Subcommittee on Energy Research and Production, the gentle lady from Tennessee, Mrs. Lloyd, would care to make a statement.

Mrs. LLOYD. Thank you, Mr. Chairman. I am pleased to be able to cochair these 2 days of hearings with you. I have had a longstanding concern about the limited access to our most advanced computers by our national laboratory and university science and engineering researchers in energy. It is clear that we as a nation do need to accelerate computer development and also that we need to establish a broader base for access to this technology. I hope that during these hearings we can clarify what the proper Federal role should be in this development.

The private sector witnesses we have here today have made tremendous strides in development of this technology on their own. Thus we must ask the question: Would government financial assistance be a help or a hindrance to these developers and will large infusions of Federal funds be necessary to assure U.S. competitiveness?

We need to take a serious look at whether the lack of accessibility to state of the art computers in the university community is taking its toll in terms of quality of training of scientists and engineers. We must also address the possibilities for additional funding for university procurement to enhance their computational ability.

According to some reports, the Japanese are only 2 years behind us in development of a fifth generation of supercomputers which would revolutionize the industry as we know it today. Given their large financial and manpower commitment to this government/industry sponsored research program, two years seems to be a very slight margin. There appear to be enormous stakes on the table here in terms of our continuing efforts to retain preeminence in this vital high technology development.

We will not only be hearing today from leading developers, but also we will hear from the major Federal users. I am pleased to welcome Dr. Robert Cooper, who has testified before my subcommittee previously on space nuclear matters, and an old friend, Dr. Al Trivelpiece, who has spoken to us often on the magnetic fusion program. Both of these gentlemen will tell us something about the vital role that their requirements can play in catalyzing advanced computer development.

I am also glad to be able to welcome back Dr. Jack Kerrebrock from NASA before he leaves to resume his post at MIT. Just as DOE, NASA is a developer as well as a user, and has been in the forefront of computer development since the early 1960's.

Gentlemen, we look forward to your testimony.

Thank you, Mr. Chairman.

Mr. FUQUA. Thank you, Mrs. Lloyd, and without objection the statements on behalf of Mr. Winn and Mr. Sensenbrenner will also be added to the record at this time.

[The opening statements of Mr. Winn and Mr. Sensenbrenner follow:]

OPENING STATEMENT BY THE HONORABLE LARRY WINN, JR.

Thank you, Mr. Chairman, for this opportunity to make a few remarks on our hearing devoted to "Computers and their Role in Energy Research."

I notice that the witnesses include industry participants beyond energy research applications. However, I have also noted through the last few years that the computer industry has become so large, so diverse, and so mobile, that it is impossible to define it briefly. Applications range from the video games we play to the war games we simulate. Over the years we have compressed compute capability from a truck load to a pocket full to do the same job. The ability to process data has been made available to anyone with $50 for a small keyboard and a T.V. set. People of all ages, backgrounds, training, and experience are active, ranging from toddlers on up in age.

To me one of the most exiting aspects of this computer revolution is the role played by industry. Early applications involved the military and space programs and have been effectively expanded and re-directed to the many other applications with which we are familiar. Throughout this transition the private sector has poured out its own funds and manpower to bring computer capability to its present status and to promise greater things in the future.

Now, however, we are advised that the present capabilities are inadequate for future military, space, and scientific needs. Terms like "artificial intelligence" and, "number crunching" become a part of our vocabulary. These needs begin to call for more expensive equipment, placing utilization in the hands of a few, primarily the government. Access to these facilities is limited, so that utilization is minimized. With this limited market potential the industrial sector is somewhat hesitant to spend necessary research and development funds.

Our witnesses should certainly be able to address these situations and many more. Some of our testimony indicates that industry appears able to devote its resources to the development of larger and more capable computers. However, we may find that there are certain rules and regulations that impede or inhibit companies from participating in advanced research and development. If such is the case, this Committee would act within its own province to alleviate or minimize those problems.

Mr. Chairman, our witnesses, both today and tomorrow, are excellent choices for the exploration of this subject. They are all very welcomed and we look forward to hearing from them.

Thank you.

STATEMENT OF THE HONORABLE F. JAMES SENSENBRENNER, JR.

Thank you, Mr. Chairman, for this opportunity to extend a welcome to our distinguished panel members.

We look forward to your testimony on the present and future role of computers in energy research.

One of the greatest testaments to the scientific and inventive capability of man has been the development of the computer. Although the computer is not a new technology, we are only beginning today to fully develop the potential of the computer. Through a partnership of academia, industry and government, this country has achieved a preeminence in the computer field. However, our preeminence in this field is being threatened.

I am particularly concerned over reports that the computer industry in this country is unable, either because of a lack of financial capital, overly restrictive antitrust laws at home, or competition from other countries, to conduct the basic research that is required for the development of the next generation of computers. These reports, if accurate, will seriously jeopardize both the economic and technological future of this country.

I look forward to your thoughts on this subject.

Mr. FUQUA. The first panel will be Mr. Larry Sumney, who is executive director of Semiconductor Research Corp., and Mr. James Hill, president and chief executive officer, Denelcor, Inc. Mr. Berlin, I hope, is in the room, if not hopefully he will be here soon, and we will go ahead, Mr. Sumney.

STATEMENTS OF LARRY W. SUMNEY, EXECUTIVE DIRECTOR, SEMICONDUCTOR RESEARCH CORP.; JAMES HILL, PRESIDENT AND CHIEF EXECUTIVE OFFICER, DENELCOR, INC.; AND JOHN F. CARLSON, VICE PRESIDENT, GOVERNMENT RELATIONS, CRAY RESEARCH, INC.

Mr. Chairman, Madam Chairperson, distinguished members of the House Subcommittees on Energy Development and Application and Energy Research and Production, ladies and gentlemen, good afternoon.

I am Larry Sumney, executive director of the Semiconductor Research Corp. This corporation has been established as a subsidiary of the Semiconductor Industry Association, and is devoted to cooperative research in semiconductors and integrated circuits through the pooling of industry resources.

I am grateful for the opportunity to appear before this committee to describe the SRC in terms of its rationale, its methods of operation, its benefits to the industries that sponsor it, as well as its relationship to supercomputer development, to academia, and to the country as a whole.

With your indulgence I have elected to use vugraphs for a major portion of my description of the SRC. As I go through and describe the research program, I would like to point out to you that the SRC is in the course of defining goals, specific research goals, for commercially available integrated circuit technology for the 1990 time period. This technology for which we believe SRC research will be of key importance, will provide the basis for progress in computers and telecommunications. This progress is directly related to the size, cost, and reliability improvements allowed by the continuing increases in density and integration level afforded by solid state science and engineering advances.

Certainly the future of supercomputers is dependent on this continuing progress. Higher speed, reduced power, new architectures, software improvements, fault tolerance, built in test, cost reduction, improved user interfaces, array processing, artificial intelligence and larger memories are all descriptive terms of needs for the next generation of supercomputers. All are dependent on improved integrated circuits.

SRC is dedicated to the research and development required to make these advances possible. If one desires supremacy in computers one must achieve supremacy in semiconductors and integrated circuits. The Japanese seem to have learned this, the Europeans have not. The United States may have almost forgotten it after very substantial original successes. Our position in the world, our defense, and the welfare of our whole society, are dependent upon recognizing this fact.

As I describe the SRC to you, you will see that it is principally focused at the present time on research, basic and some applied. However, the SRC is considering extending its mission into areas other than pure research. In particular, the industry is taking a broader view of cooperative efforts; it is requiring focus on manufacturing automation and productivity improvements.

Combining the research work of the universities with demonstration manufacturing facilities, allows us to develop products that under normal circumstances would not be developed until much later in time. These other mission areas are currently in the planning process and I am not prepared to describe them for you today, but they do extend beyond the basic research which I will briefly sketch for you at this time.

The concept of the SRC grew out of extensive planning by the semicondutor industry to develop an organization that would provide knowledgeable and professional planning of a broad industrywide research program. The SRC exists through sponsorship by its member companies through a fee structure that is based on a company's sales; if it sells integrated circuits, purchases—if it buys them to use in equipment, or semiconductor related sales; if it is in the business of selling equipment or materials to the semiconductor industry.

The research that we do is coordinated through an industry SRC technical advisory board. The research programs principally at this time are in the university community, and they are contracts rather than grants.

We are very active. One of our major roles is trying to achieve an effective and efficient transferring of the research results back into our sponsoring industries.

I would like to point out that the SRC is a new organization. We were incorporated in February 1982. I was hired in May 1982, and at this point in time we have essentially contracted—this says two centers and and five elements, which are essentially at three centers and 42 project awards at this point in the university community, and I will describe those briefly for you as I go through the presentation.

Our membership in the SRC at this point involves a cross section of companies. They range from large companies, such as Motorola and IBM, to medium size companies, such as Hewlett-Packard, to small companies, such as Silicon Systems and Monolithic Memories.

Another cut—they range from semiconductor companies, such as Advanced Micro Devices, Intel Corp., to systems houses such as IBM, Digital Equipment Corp., to equipment manufacturers, such as Eaton, to material manufacturers, such as Monsanto, which produces silicon for the semiconductor industry.

We are continuing to try to expand our membership. Some of the companies that have expressed an interest in joining are shown here. Candidates that we are beginning to have discussions with are shown here and new candidates for solicitation are shown in the last column here.

Early on in the formation of our research program, we identified eight major thrust areas that we felt were of key importance to the semiconductor industry—design automation, semiconductor materials, system and component interaction, which is an architecture system, reliability, quality assurance and tests, small device structures, research into new methods of device fabrication, research into the manufacturing processes, research into packaging and higher level or off chip interconnection.

Given the eight major research thrusts, we identified a program structure. Each of our major thrusts is shown here only on this chart. They are grouped, three, two, three, with these three sort of grouping into a general area of microstructure sciences; these two sort of grouping into systems and design; these three into production and engineering.

The general thought is that we will have a university campus identified as the center of excellence doing SRC sponsored research in each of these eight areas. Those centers will be supported by a number of smaller programs at universities across the country. Typically the centers will be funded at $1 million a year or above; the projects over here of which we now have 42 will be funded at nominally at $100,000 per year, and typically will involve one faculty member and one or two graduate students.

Over time these centers here, one will be identified as a lead center in small structured sciences; one will be identified as a lead center in systems and design; and one of these will be identified as a lead center in production and engineering.

For program flexibility, I have described what I have just mentioned. We have a new thrust category. Those are areas of research concentration that are generated by need or opportunity. Currently this week we are sponsoring a workshop in North Carolina in the area of the five semiconductor materials trying to see what the SRC can effectively do in the arena in terms of generating a research program to be responsive to the industry needs.

We are taking into consideration research that is ongoing, particularly that which is sponsored by the Government.

It is useful to take that organization or structure, put a time line on it, just to show you where we have gone since May until this time in 1 year. As I indicated, we have negotiated and signed 42 projects at $100,000 a year at universities across the country. We have identified our computer aided design centers of excellence. We have two campuses. In this case, one at Carnegie-Mellon University, and one at Cal-Berkeley. We have identified our microstructure science center of excellence. That is at Cornell University. These show those yet to be identified. The identification of lead centers is an evolutionary thing that will happen in 1984 when the first out of this set will be identified.

One of the key mechanisms that we want to do to tie the organization together is to establish a computer network, which we are currently doing. That will tie the centers of excellence, the supporting research activities and the other universities, plus the lead centers and the SRC together via a network that most likely will involve the already existing computer science network that exists today in the university community and with some companies.

I am not going through this. It is simply meant to point out the three major areas that I have identified are the eight thrusts, as shown here. This then ties in the contracts that we have awarded and sort of shows with respect to the technical area, the distribution of contract awards.

You note that we have only two contracts in the architecture area, only two contracts in the areas of reliability, quality assurance and test, and what that tells us is, since many of these contracts were proposals in the university community, it is in these areas if we are going to do meaningful research, we need to nurture a resource in the university community to move forward in these areas.

Distribution of SRC funding relevant to present activity. Arbitrarily we have identified Cornell, Berkeley, Sandia, MIT, Illinois, and Standford for those schools that we fund as first tier schools in semiconductor R&D, and we have currently put about $4.8 million annual rate of funding into those schools to support SRC contracts.

In tier two schools, we have put approximately the same amount of money for a total obligation at this time at annual rate of $8.8 million.

One of our major goals is to try to build up these programs that show promise so that many of these schools that are currently in tier 2 categories in this area of research will be moved over into the tier 1 area and that resource can be expanded upon.

Regionally, our distribution of SRC funding looks like this at this time. It is geographically pretty well distributed, and I don't think there is any major area that doesn't have at least some degree of SRC funding at this point. This was not a major factor in our awards, it just turned out that there is a reasonably good degree of geographical distribution of funding.

Another major objective in addition to doing good research, is trying to improve the quantity and quality of graduate students available to the industry. We are not going to do that through fellowships. We are leaving that to the sponsoring firms to administer, but on the other hand, one of the reasons why we are executing our research program in the university community is to try to enhance U.S. capability to train graduate students, and we are doing that by structuring research programs that are interesting and that have funding continuity, and that way we hope to attract undergraduates to graduate school and keep them there, and better attract competent professors and instructors to SRC supported graduate schools. So our educational objectives are important, but we are sort of approaching them indirectly.

I am not going to go through the fee schedule in detail, but essentially companies that are producers of integrated circuits pay on their worldwide IC sales. Companies that are users of integrated circuits pay on their worldwide IC purchases. And as a rough estimate, their assessment is approximately a tenth of 1 percent of those figures.

As the semiconductor industry examined world trends in semiconductor research, they saw the efforts in Japan in nationalization of industry and government funding of research to meet economic and political goals.

There are some significant startup efforts in China.

England has started a program called VHPIC, which is a military oriented program, standing for very high performance integrated circuits. They also have the INMOS Co., of which one of those fabrication facilities is in the United States.

The United States has the VHSIC program which I started in the Department of Defense prior to taking on this responsibility. That is a government joint effort. It was not established to meet industry needs in this arena, but rather it was established to direct industry concerns toward the unique needs of the military. Those unique needs involved higher speed for signal processing, radiation hardening, perhaps density quicker than the industry demands. It is really a lead on the industry, but many of them are unique, such as the radiation hardness and signal processing.

In conclusion, the semiconductor industry saw a fragmentation of research resources in companies that resulted in a below critical mass support to many of the areas that they felt were of key importance. They saw focus of government support of semiconductors in computer research in foreign countries without corresponding U.S. response and a rapidly eroding preeminence of this industry in the United States.

They saw lack of support in semiconductor universities of faculty and graduate students and that perhaps is more accurately stated by saying that the support was there but the university faculty had to spend a tremendous amount of time going out and organizing it and structuring the support and finding it. What we are trying to do here is to relieve that burden, to provide some continuity and promise of funding to good research programs.

The solution that the industry came up with was the SRC, where the cost of generic research is to be shared across the industry to gain some type of leverage. Now, if we went over the fee structure in detail, no company can contribute more than 10 percent to our approved budget, which is $12 million this year. So no company can be assessed for more than $1.2 million regardless of what their sales are. Assuming we had 10 of those big companies, each would gain some leverage of about 10 to 1. When you look at the smaller companies that can come in for $50,000, which is our minimum assessment, the leverage factor is significantly greater.

Alternatives, mergers to massive entities, government sponsorship of programs, and resulting control and direction, guard their unique needs. Business erosions are things the semiconductor industry felt were not attractive. SRC was formed to be an alternative to those potential solutions.

[The prepared statement of Mr. Sumney follows:]

SEMICONDUCTOR RESEARCH CORPORATION

SRC
COOPERATIVE
RESEARCH

June 14, 1983

The attached is for testimony at a joint hearing before the Subcommittee on Energy Development and Applications and the Subcommittee on Energy Research and Production on June 14, 1983, from 2:00 p.m. to 4:00 p.m. in Room 2318 of the Rayburn House Office Building.

I am Larry W. Sumney, Executive Director of the Semiconductor Research Corporation, a subsidiary of the SIA devoted to cooperative research through the pooling of industry resources.

I am grateful for the opportunity to appear before this committee to describe the SRC in terms of its rationale, its modus operandi, its benefits to the industry that sponsor it, as well as its relationships to supercomputer development, to academia, and to the country as a whole.

300 PARK DRIVE, SUITE 215 • P.O. BOX 12053 • RESEARCH TRIANGLE PARK, N.C. 27709 • (919) 549-9333

THE SEMICONDUCTOR RESEARCH CORPORATION

INTRODUCTION

The Semiconductor Research Corporation (SRC) is a new organization that is implementing an innovative form of industry-university cooperation in research related to semiconductor devices. The SRC's goals are to plan, promote, coordinate, conduct, and sponsor research that will result in (1) new knowledge of semiconductor materials and phenomena, and of related scientific and engineering subjects that are required for the useful application of semiconductors; (2) the development of new and more efficient design and manufacturing technologies for semiconductor devices; and (3) an increase in the number of scientists and engineers that are proficient in research, development, and manufacture of semiconductor devices. An equally important SRC task is to efficiently communicate the results of its research to its industrial members. Since the SRC has not yet completed the first year of its operations, this description is a snapshot of an organization and program that is far from maturity.

BACKGROUND

The SRC is a natural derivative of the development of semiconductor technology in the U.S. The generic technology on which this industry is based was spawned by the Bell Telephone Laboratories in the fifteen years following the invention of the transistor with a major infusion of funding from the defense establishment for research in both industrial and university laboratories. The growth of the semiconductor device industry was so rapid that by the mid-1960's developments within the industry were outpacing the ancillary research community. Both Bell and the Department of Defense (DOD) refocused their efforts to the application of the fruits of their earlier research rather than the continued advancement of the generic technology. Support of academic research shifted in emphasis to longer range, more speculative areas such as compound semiconductors. Research directed to both the near-term and long-range needs of the mainstream semiconductor industry was largely confined to the laboratories of the companies producing the devices. Both merchant producers and the growing number of captive semiconductor producers were involved. These circumstances had negative impacts. First, the competitive companies were reluctant to share research results that might provide a competitive advantage. This slowed, but did not stop the diffusion of new results. Second, the research efforts became highly redundant. Each company found it necessary to apply its efforts to very similar topics in order to avoid being left behind in a critical, rapidly developing technology. Gradually, research became focused on solution of near-term problems, and the generic technology for the longer range future was neglected. Third, and most important, the structure of the industry and the economy resulted in an erosion of the resources available for research, and it was necessary for each company to apply these resources to current product development in order to remain competitive.

It is important to note that research in compound semiconductors is an important element of the generic technology for future semiconductor products. However, it represents just one of the areas that need to be addressed. It is just as possible that new generations of capabilities may be based on the evolution of silicon technology into radically new device structures as it is that new but less developed material systems, such as the III-V semiconductors, will provide these capabilities. Each candidate technology must be addressed in a comprehensive generic technology at a level commensurate with its perceived role.

In late 1981, executives of major U.S. companies that produce and/or use semiconductor products recognized that this erosion of the generic technology base coupled with government financed efforts in other countries constituted an important competitive threat to their industry. Recognizing that a government-based response to this threat was unlikely, they decided to undertake a cooperative industry-initiated response. After considerable discussion, a concept emerged that has become the SRC. This concept is that SRC will:

- Be sponsored by member companies through fees based on semiconductor sales or usage, or semiconductor-related sales.
- Provide knowledgeable and professional planning of an industry-wide research program.
- Contract for research to be coordinated by an industry/SRC technical advisory board.
- Play an active role in monitoring research progress and transferring results back to industry.
- Establish both
    - University centers of excellence with long-term, major, thrust areas.
    - Additional research activities based on shorter term, more precisely defined objectives.
- Result in
    - Establishment of a major, focused, generic research program supportive of the needs of the semiconductor industry.

- More relevant graduate education and a larger supply of graduate students.
- Improved communication within the U.S. semiconductor community.

Subsidiary research objectives that are consistent with this concept include objectives such as to:

- Provide a clearer view of directions, limits, opportunities, and problems in generic technologies.
- Decrease fragmentation and redundancy in U.S. semiconductor research.
- Establish above-threshold research efforts for critical areas that are beyond the resources of many cooperating companies.
- Enhance the image of the industry in order to attract needed talent.
- Increase industry-university ties.
- Provide a scientific and technical information base for industry development efforts.
- Attract highly qualified students to graduate programs in semiconductor research.
- Help universities to attract and retain competent faculty in semiconductor related areas.

As part of this set of objectives, it was agreed that the SRC program should encompass universities that are not presently leaders in semiconductor research in order to broaden the institutional base, while increasing the level of effort in universities already active in this technology.

HISTORY AND ORGANIZATION

By early 1982 the SRC concept was sufficiently advanced to where, under the aegis of the Semiconductor Industries Association, the SRC was incorporated and by-laws were written. In May 1982, Larry W. Sumney, former director of the DOD Very High Speed Integrated Circuit Program, was hired as Executive Director of the SRC. In September 1982, offices of the SRC were established in the Research Triangle Park of North Carolina, initial staff was hired, and efforts to establish the research program were initiated. In November, eight universities were selected for award of the first research contracts and 166 proposals were received from 63 universities. By June of this year, over thirty universities are participating in the SRC research program.

The SRC exists as a cooperative effort of companies that produce or use semiconductor devices or are major suppliers to the semiconductor industry. Present membership includes:

Advanced Micro Devices, Inc.
Control Data Corporation
Digital Equipment Corporation
E-Systems
Eaton Corporation
General Electric
General Instrument Corporation
Harris Corporation
Hewlett-Packard Company
Honeywell, Incorporated
IBM Corporation
Intel Corporation
Monolithic Memories, Incorporated
Monsanto Company
Motorola, Incorporated
National Semiconductor Corporation
SEMI (Chapter)
Silicon Systems, Incorporated
Westinghouse Electric Corporation

By the end of 1983 over two dozen companies are expected to be members of the SRC and supporting its research activities on the basis of an established fee schedule.

The SRC reports to a Board of Directors (BOD) which establishes major policy for it, approves major contract awards, approves the appointment of Technical Advisory Board members and key members of the professional staff, and oversees the development and operations of the organization.

The SRC TAB has an active role in the evolution of the research program of the SRC participating in the evolution of the research strategy and the evaluation of proposals. It will provide an important conduit for the flow of information from industry to SRC research contracts, and for the flow of research results from the SRC to industry.

The SRC staff, under the direction of the BOD and with the advice of the TAB, formulates and implements the research program of the SRC. It is intended that the permanent staff remain small so that the active participation of the BOD, TAB, and industry assignees will continue to be necessary in order to carry out, manage, and direct the SRC program. In addition, it is planned that each contractor will have an industry individual identified as an industry resource contact to provide assistance and information when required.

RESEARCH PROGRAM PLAN

Before the SRC research program was initiated, a strategy for the program was adopted that identifies the major areas of interest. The major areas identified in this strategy are:

- Microstructure Sciences
    - Materials, Phenomena, and Device Physics
    - Microsciences
    - Device Fabrication Technologies

- Systems and Design
    - Design Automation
    - System - Component Interactions

- Production and Engineering
    - Reliability, Quality Assurance, and Testing
    - Packaging
    - Manufacturing

A three-tier layer of research activities is envisioned consisting of individual project awards, centers of excellence, and lead centers. The first level is intended to comprise those research activities centered around one or several faculty members and their students awarded in a competitive process based on solicitation of proposals for innovative research. The implementation of the first of these solicitations will be described in the next section. At the second level, centers of excellence are being established both through negotiation and competition. These centers will comprise team efforts with themes that can correspond to major areas in the strategy or other mutually acceptable topics that are deemed to be of significant importance. An effort is being made to supplement center themes with research vehicles that provide additional focus to the efforts. The third tier of planned research activity is identified as a lead center and conceptually provides a framework for amalgamating related efforts and facilities of multiple institutions into a structure of coordinated research. The implementation of lead centers will evolve from the present project-center activities at some time in the future.

In addition to the three-tier structure of research described above, an additional program component referred to as a "new thrust" is provided. This is a mechanism for implementation of focused research efforts that respond to needs, opportunities, or technological threats beyond the outlined research strategy. Such efforts will be initiated either through competitive solicitations or from unsolicited proposals. Examples might include areas such as compound semiconductors, high power IC's, novel architectures, electro-optical IC's, or precision processing.

RELATIONSHIP TO COMPUTER DEVELOPMENT

In its research program, the SRC is in the course of defining goals for a commercially available integrated circuit technology for the 1990 time period. This technology--for which we believe SRC research will be of key importance--will provide the basis for progress in computers and communications. This progress is directly related to the size, cost, and reliability improvements allowed by the continuing increases in density and integration level afforded by solid state science and engineering. Certainly, the future of supercomputers is dependent on this continuing progress.

Higher speed, reduced power, new architectures, software improvements, fault tolerance, cost reduction, improved user interfaces, array processing, artificial intelligence and large memories are all descriptive terms of needs for the next generation of supercomputers. All are dependent on improved integrated circuits. The SRC is dedicated to the research and development required to make these advances possible. If one desires supremacy in computers, he must achieve supremacy in semiconductors and integrated circuits. The Japanese seem to have learned this, the Europeans have not, and the U.S. has almost forgotten it after original successes. Our position in the world, our defense, and the welfare of our whole society are dependent on recognizing this fact.

PROGRAM INITIATION

The initial implementation of the SRC research program is based on a broad solicitation for research proposals that was distributed in September and October of last year and on several unsolicited proposals. As noted, the competitive solicitation resulted in 166 proposals that were evaluated and ranked in the three major categories of the research strategy. This process resulted in a decision to make awards of approximately forty contracts to twenty-six universities. In addition, thirty-seven proposals from two universities, RPI and MIT, and one university group, the Microelectronics Center of North Carolina, were removed from the competitive evaluation and are serving as a basis for the establishment of three well-focussed research programs. In parallel with this broad solicitation, centers-of-excellence have been established at Cornell University in the area of microscience and technology and at the University of California/Berkeley and Carnegie-Mellon University in computer aided design.

SUMMARY

The SRC with a 1983 budget of approximately $11 million is rapidly establishing a structured research program that responds to the generic technology needs of the semiconductor industry. Its initial research efforts have focused on addressing nearer-term problems; however, in order to achieve balance and maximize the long-term return-on-investment to its members, the SRC will address the full range of science and technology that will significantly contribute to the advancement of semiconductor technology both in this decade and on to the turn of the century. With the support of a growing number of industrial participants, this research will become a major factor in establishing the technology for the information age and will assure the U.S. industry a position of continued leadership in the semiconductor field as well as in the myriad application areas that depend on advanced semiconductor devices.

Mr. HILL. Mr. Chairman, Madam Chairperson, distinguished members of the House Subcommittees on Energy Development and Applications and Energy Research and Production, ladies and gentlemen, good afternoon.

I am James A. Hill, president and chief executive officer of Denelcor, Inc., a publicly held Aurora, Colo. company which designs and produces large scale, high performance computers for scientific and technical application. On behalf of Denelcor's computer scientists, engineers, and business associates, I would like to express our appreciation of your interest and foresight in conducting these timely hearings on significant Department of Energy matters. It is a pleasure to have this opportunity to testify before you on one of the most important issues facing our Nation today: America's continued preeminence in supercomputer technology.

The United States has been the leader in large scale computing systems since they were first created in the postwar years of the 1950's. Development accelerated during the 1960's when our Nation gave high priority to scientific achievement as we set out to land a man upon the moon. During the 1970's, however, support for scientific and technological advancement declined in the United States. Efforts intensified in other countries, such as Great Britain, West Germany, France, and Japan, to achieve predominence in high speed computers.

Today, as you are well aware, the Japanese have embarked on an ambitious program to surpass America in the area of supercomputers. To date, most discussion has centered on the so-called fifth generation project, a joint government, business, and education program, aimed at establishing an artificial intelligence industry by the mid-1990's. Less attention, however, has been focused on Japan's other supercomputer project—the one aimed at developing an ultra high speed computing system for scientific and technological application. The stated Japanese objective is to produce a supercomputer by 1989 which would be 1,000 times as fast and as powerful as any scientific computing system currently on the market.

While some observers have dismissed the Japanese super speed computer project as unrealistic, we at Denelcor believe that development of such a system is eminently feasible, particularly since we understand that the Japanese plan envisions the creation of a parallel processing architecture, similar to the high performance architecture we have been successfully developing for the past 7 years.

As computer engineers and businessmen, we also recognize that development of an ultra high speed computer system by Japan or any other country represents a serious threat to the United States—to our national security, our economic viability, and our scientific and technological leadership in the world. The prospect of U.S. dependence on foreign scientific supercomputer technology has far more serious ramifications than any previous U.S. dependency on OPEC.

This afternoon I would like to address this pressing issue and direct my brief remarks to the strategic, economic and scientific importance of maintaining American leadership in the development, manufacture, and use of large scale high performance scientific computing systems, and the vital roles that government can play in assuring American supercomputer preeminence in the future.

Scientific supercomputers are the most advanced computing systems in the world. Our so-called "number crunchers" are fundamentally different from the more familiar data processing computers. They have different architectures and are significantly faster than the general purpose business machines.

Scientific supercomputers are specifically designed to handle the most computationally intensive problems, such as those encountered in nuclear physics. They are problems which involve constantly changing variables, many unknowns, and often require millions of different instructions and billions of different computer calculations for solutions.

To help you gain a better understanding of the complexity and importance of scientific supercomputers, let me outline the role they play, for example, in weather forecasting, something that affects all of us on a daily basis.

To make an accurate weather forecast, a meteorologist attempts to take into account every available piece of information he can obtain on weather around the globe. He receives reports from worldwide observation stations which are constantly being revised and updated, and he reviews this information in light of his own knowledge of the dynamics and basic physical laws of atmosphere.

As he becomes more sophisticated, he extends his understanding of basic fluid dynamics to include geophysical rotational effects and specific problems of turbulence. Later, he begins considering other physical phenomenon, such as evaporation and condensation of water, solar heating, infrared cooling and surface friction and drag. Before too long, he is adding three dimensional cloud physics, with variable droplet size and charge spectra, into his computations while still attempting to make the most accurate forecast he can manage by a time critical deadline.

How can he possibly assimilate all that information? He can't, unless he has use of a scientific supercomputers, a computing system powerful enough to process billions of pieces of information and fast enough to make all those millions of calculations, timely and worthwhile.

Thus, speed and performance become critical to a scientist's specific solutions to complex scientific problems:

Solutions to problems of national security, including ballistic missile trajectories, defense systems, preparedness, nuclear weapons development and simulation of nonatmospheric intelligence data analysis and cryptography;

Solutions to complex problems of economic and commercial importance to the United States, such as the aerodynamic design and simulation testing of a jet aircraft before it is built, the discovery of petroleum and mineral deposits from two and possibly three dimensional seismic data, the investigation of our oceans without disruption to the ecology, the solution to questions concerning electronic circuitry, electric power distribution systems, structural engineering models, robotics, and many other problems of commercial and economic significance to this country.

And the solution of complex problems of safety and health, including the development of safety systems to prevent nuclear reactor accidents, the creation of micro-biotic organisms to contain and eliminate hazardous wastes, the development of future generations of cat scanners to assist in medical diagnosis, and the solution of hundreds of biomedical, genetic, molecular and contagious disease problems.

In short, scientific supercomputers have become tremendously valuable in solving complex problems vital to our national security, our economic and commercial leadership in the free world and the health, safety and well being of our citizenry. It would be distressing, to say the least, if America lost its supercomputer leadership.

At Denelcor, we have no intention of giving up our leadership position to the Japanese or anyone else. We may not be as well known as some others in the industry, but I am proud to report to this distinguished panel that recently we made what may be the most substantial breakthrough in the history of supercomputers technology.

When the Japanese embarked on the so-called national superspeed computer project 18 months ago, we were already 4 years ahead of them in the parallel processing architecture and were in the process of manufacturing our first parallel processor computing system for the U.S. Army Ballistics Research Laboratory at Aberdeen, Md., under the dedicated sponsorship of Dr. Eichelberger. Installed last fall, our four processor "HEP" system is now recognized as the world's fastest scalar computer, with a capacity to process 40 million instructions per second.

In developing the HEP system, Denelcor effectively overcame the traditional problem encountered in multiple unit systems, namely the loss of a linear increase in speed and performance when each processor unit is added to the system. Our successful application of parallel process architecture has enabled us to put multiple units together and achieve a direct linear increase in speed and performance. A 16-unit HEP today makes Denelcor's system significantly faster and more powerful than any existing supercomputers of its type on the market.

Recently, we completed the basic design study of Denelcor's HEP 2 system, and in that process we made a monumental breakthrough. Our new system, in its largest configurations, will be capable of processing in excess of 12 billion instructions per second. That is about 500 times faster and more powerful than any currently available supercomputers, and about half the projected level that the Japanese are striving to reach by the end of the decade. Denelcor will have its HEP 2 system on the market by early 1986, and expects to continue the natural evolution of its architecture to exceed the stated Japanese national goal.

We achieved these results largely through our own internally financed research and development activities and without government assistance or funding. We believe there is tremendous potential for the use of our supercomputers systems.

This brings me to the most important remarks I would like to make today—comments concerning the role government can play maintaining American preeminence in supercomputers technology. As a businessman and computer developer, I am less concerned about the danger of losing our leadershp in the design of supercomputers than I am about the dangers of losing our leadership in supercomputers usage.

Small companies like Denelcor are in an excellent position to develop ideas—new concepts and innovative designs—through the private development. But if our good ideas are to be of any lasting value, they must be exploited fully, and full exploitation is generally beyond the means or even the primary interests of companies our size.

During the 1950's and 1960's the government placed a high priority on the exploitation of scientific ideas, and we believe this commitment should be renewed in the 1980's. Let us look beyond our present interests in production computers, essential as they may be, and begin preparing for the future.

Today, the laboratories are largely engaged and focus their funding on steps that are repetitive that have a tendency in their minds to be productive as opposed to being experimental in research.

The most effective way government can fully exploit our present supercomputers preeminence is to put our machines to use in solving the immensely complex problems our nation and world face today. In short, I urge government to renew the national commitment to research and development.

I urge all of you in government to provide the leadership our nation needs to keep America in the forefront of scientific and technological advancement, and to encourage and train new generations of scientists and engineers who can maximize the problem solving capabilities of our advanced scientific computing systems.

Specifically, we at Denelcor urge you to:

One, take the leadership role in to formulating a national plan to achieve specific scientific and technological objectives over the next one or two decades, and grant the Department of Energy responsibility for directing a coordinated, yet distinctly independent, government, business and educational effort to implement our national plan. If we in business know what our national priorities are, we can achieve them.

I am not trying to propose a national funding plan. We do need a leadership goal, the industry needs a target stated by a coordinated interagency group. I believe if it is stimulated properly that private industry will know what to do with that and can move and take the attendent risks.

Two, increase funding for research and development projects, enhanced supercomputers capacity and staffing at the national laboratories of the Department of Energy, the research facilities of the Department of Defense, the National Aeronautics and Space Administration, National Science Foundation and in the intelligence community. These in general are in our view, intelligent, smart users of computers. Government support of our Nation's most talented, capable, and sophisticated supercomputer scientists is the most expedient and least expensive way to provide support to complex problem solutions through supercomputer research.

Three, open your doors and provide supercomputer access to other interested scientists and engineers by establishing research, development and problem solving networks reaching out from the national laboratories to regional government facilities and beyond to private research agencies and our nation's universities.

Four, encourage the development of component and packaging technology for U.S. supercomputers of the 1980's and 1990's to alleviate our present and projected dependence on foreign suppliers of computer components. There are elements of components and packaging today that are at hand and in the control of the Japanese companies, and we have recently been refused access to some of this technology, and the statement we got was "come back in 2 years, we will make it available to you after our indigenous manufacturers have been using it." These things are very scary.

Five, stimulate development of software for high performance computing systems, including the design of algorithms, numerical analysis and computational mathematics to maximize the performance potential of new generation supercomputers. Many of the research areas are beyond the reach of a company our size and I believe this is an area especially in the software research where government can participate through the universities and laboratories.

Six, spark the imagination of young scientists and engineers by providing opportunities for training and real time use of supercomputers through the establishment of fellowships, loans, and special work programs between government agencies and universities.

By taking these actions today, you in government will secure American supercomputer leadership in the years ahead. We in industry can maintain America's lead in the design and manufacture of super scientific computing systems if you in government can provide the lead in putting our supercomputers to work in scientific and technological research and development. Together we can solve the complex problems of our times.

Thank you very much. I will be happy to answer questions.

[The prepared statement of Mr. Hill follows:]

TESTIMONY BY

JAMES A. HILL
PRESIDENT AND CHIEF EXECUTIVE OFFICER
DENELCOR, INC.

MR. CHAIRMAN, MADAM CHAIRPERSON, DISTINGUISHED MEMBERS OF THE HOUSE SUBCOMMITTEES ON ENERGY DEVELOPMENT AND APPLICATIONS AND ENERGY RESEARCH AND PRODUCTION, LADIES AND GENTLEMEN, GOOD AFTERNOON.

I AM JAMES A. HILL, PRESIDENT AND CHIEF EXECUTIVE OFFICER OF DENELCOR, INC., A PUBLICALLY-HELD AURORA, COLORADO COMPANY WHICH DESIGNS AND PRODUCES LARGE-SCALE, HIGH PERFORMANCE COMPUTERS FOR SCIENTIFIC AND TECHNICAL APPLICATION. ON BEHALF OF DENELCOR'S COMPUTER SCIENTISTS, ENGINEERS AND BUSINESS ASSOCIATES, I WOULD LIKE TO EXPRESS OUR APPRECIATION OF YOUR INTEREST AND FORESIGHT IN CONDUCTING THESE TIMELY HEARINGS ON SIGNIFICANT DEPARTMENT OF ENERGY MATTERS. IT IS A PLEASURE TO HAVE THIS OPPORTUNITY TO TESTIFY BEFORE YOU ON ONE OF THE MOST IMPORTANT ISSUES FACING OUR NATION TODAY: AMERICA'S CONTINUED PRE-EMINENCE IN SUPER COMPUTER TECHNOLOGY.

* * *

THE UNITED STATES HAS BEEN THE LEADER IN LARGE SCALE COMPUTING SYSTEMS SINCE THEY WERE FIRST CREATED IN THE POST WAR YEARS OF THE 1950'S. DEVELOPMENT ACCELERATED DURING THE 1960'S WHEN OUR NATION GAVE HIGH PRIORITY TO SCIENTIFIC ACHIEVEMENT AS WE SET OUT TO LAND A MAN UPON THE MOON. DURING THE 1970'S, HOWEVER, SUPPORT FOR SCIENTIFIC AND TECHNOLOGICAL ADVANCEMENT DECLINED IN THE U.S. EFFORTS INTENSIFIED IN OTHER COUNTRIES SUCH AS GREAT BRITAIN, WEST GERMANY, FRANCE, AND JAPAN TO ACHIEVE PREDOMINENCE IN HIGH SPEED COMPUTERS.

TODAY, AS YOU ARE WELL AWARE, THE JAPANESE HAVE EMBARKED ON AN AMBITIOUS PROGRAM TO SURPASS AMERICA IN THE AREA OF SUPERCOMPUTERS. TO DATE, MOST DISCUSSION HAS CENTERED ON THE SO CALLED "FIFTH GENERATION" PROJECT, A JOINT GOVERNMENT, BUSINESS AND EDUCATION PROGRAM, AIMED AT ESTABLISHING AN ARTIFICIAL INTELLIGENCE INDUSTRY BY THE MID-1990'S. LESS ATTENTION, HOWEVER, HAS BEEN FOCUSED ON JAPAN'S OTHER SUPERCOMPUTER PROJECT -- THE ONE AIMED AT DEVELOPING AN ULTRA HIGH SPEED COMPUTING SYSTEM FOR SCIENTIFIC AND TECHNOLOGICAL APPLICATION. THE STATED JAPANESE OBJECTIVE IS TO PRODUCE A SUPERCOMPUTER BY 1989 WHICH WOULD BE 1000 TIMES AS FAST AND AS POWERFUL AS ANY SCIENTIFIC COMPUTING SYSTEM CURRENTLY ON THE MARKET.

WHILE SOME OBSERVERS HAVE DISMISSED THE JAPANESE SUPER-SPEED COMPUTER PROJECT AS UNREALISTIC, WE AT DENELCOR BELIEVE THAT DEVELOPMENT OF SUCH A SYSTEM IS EMINENTLY FEASIBLE, PARTICULARLY SINCE WE UNDERSTAND THAT THE JAPANESE PLAN ENVISIONS THE CREATION OF A PARALLEL PROCESSING ARCHITECTURE, SIMILAR TO THE HIGH-PERFORMANCE ARCHITECTURE WE HAVE BEEN SUCCESSFULLY DEVELOPING FOR THE PAST SEVEN YEARS.

As computer engineers and businessmen, we also recognize that development of an ultra high speed computer system by Japan or any other country represents a serious threat to the United States - to our national security, our economic viability and our scientific and technological leadership in the world. The prospect of U.S. dependence on foreign scientific supercomputer technology has far more serious ramifications than any previous U.S. dependency on UPEC.

This afternoon I would like to address this pressing issue and direct my brief remarks to the strategic, economic and scientific importance of maintaining American leadership in the development, manufacture and use of large scale high performance scientific computing systems, and the vital roles that government can play in assuring American supercomputer pre-eminence in the future.

* * *

Scientific supercomputers are the most advanced computing systems in the world. Our so called "Number Crunchers" are fundamentally different from the more familiar data processing computers. They have different architectures and are significantly faster than the general purpose business machines.

SCIENTIFIC SUPERCOMPUTERS ARE SPECIFICALLY DESIGNED TO HANDLE THE MOST COMPUTATIONALLY INTENSIVE PROBLEMS, SUCH AS THOSE ENCOUNTERED IN NUCLEAR PHYSICS. THEY ARE PROBLEMS WHICH INVOLVE CONSTANTLY CHANGING VARIABLES, MANY UNKNOWNS, AND OFTEN REQUIRE MILLIONS OF DIFFERENT INSTRUCTIONS AND BILLIONS OF DIFFERENT COMPUTER CALCULATIONS FOR SOLUTIONS.

TO HELP YOU GAIN A BETTER UNDERSTANDING OF THE COMPLEXITY AND IMPORTANCE OF SCIENTIFIC SUPERCOMPUTERS, LET ME OUTLINE THE ROLE THEY PLAY, FOR EXAMPLE, IN WEATHER FORECASTING, SOMETHING THAT AFFECTS ALL OF US ON A DAILY BASIS.

TO MAKE AN ACCURATE WEATHER FORECAST, A METEOROLOGIST ATTEMPTS TO TAKE INTO ACCOUNT EVERY AVAILABLE PIECE OF INFORMATION HE CAN OBTAIN ON WEATHER AROUND THE GLOBE. HE RECEIVES REPORTS FROM WORLD-WIDE OBSERVATION STATIONS WHICH ARE CONSTANTLY BEING REVISED AND UPDATED, AND HE REVIEWS THIS INFORMATION IN LIGHT OF HIS OWN KNOWLEDGE OF THE DYNAMICS AND BASIC PHYSICAL LAWS OF ATMOSPHERE. AS HE BECOMES MORE SOPHISTICATED, HE EXTENDS HIS UNDERSTANDING OF BASIC FLUID DYNAMICS TO INCLUDE GEOPHYSICAL ROTATIONAL EFFECTS AND SPECIFIC PROBLEMS OF TURBULENCE LATER, HE BEGINS CONSIDERING OTHER PHYSICAL PHENOMENON, SUCH AS EVAPORATION AND CONDENSATION OF WATER, SOLAR HEATING, INFRA-RED COOLING AND SURFACE FRICTION AND DRAG. BEFORE TOO LONG, HE'S ADDING THREE-DIMENSIONAL CLOUD PHYSICS, WITH VARIABLE DROPLET SIZE AND CHARGE SPECTRA, INTO HIS COMPUTATIONS WHILE STILL ATTEMPTING TO MAKE THE MOST ACCURATE FORECAST HE CAN MANAGE BY A TIME-CRITICAL DEADLINE.

HOW CAN HE POSSIBLY ASSIMILATE ALL THAT INFORMATION? HE CAN'T -- UNLESS HE HAS USE OF A SCIENTIFIC SUPERCOMPUTER, A COMPUTING SYSTEM POWERFUL ENOUGH TO PROCESS BILLIONS OF PIECES OF INFORMATION AND FAST ENOUGH TO MAKE ALL THOSE MILLIONS OF CALCULATIONS, TIMELY AND WORTH WHILE.

THUS SPEED AND PERFORMANCE BECOME CRITICAL TO A SCIENTIST'S SPECIFIC SOLUTIONS TO COMPLEX SCIENTIFIC PROBLEMS:

o SOLUTIONS TO PROBLEMS OF NATIONAL SECURITY, INCLUDING BALLISTIC MISSILE TRAJECTORIES, DEFENSE SYSTEMS, PREPAREDNESS, NUCLEAR WEAPONS DEVELOPMENT AND SIMULATION OF NON-ATMOSPHERIC INTELLIGENCE DATA ANALYSIS AND CRYPTOGRAPHY;

o SOLUTIONS TO COMPLEX PROBLEMS OF ECONOMIC AND COMMERCIAL IMPORTANCE TO THE UNITED STATES, SUCH AS THE AERODYNAMIC DESIGN AND SIMULATION TESTING OF A JET AIRCRAFT BEFORE IT IS BUILT, THE DISCOVERY OF PETROLEUM AND MINERAL DEPOSITS FROM TWO - AND POSSIBLY THREE-DIMENSIONAL SEISMIC DATA, THE INVESTIGATION OF OUR OCEANS WITHOUT DISRUPTION TO THE ECOLOGY, THE SOLUTION TO QUESTIONS CONCERNING ELECTRONIC CIRCUITRY, ELECTRIC POWER DISTRIBUTION SYSTEMS, STRUCTURAL ENGINEERING MODELS, ROBOTICS, AND MANY OTHER PROBLEMS OF COMMERCIAL AND ECONOMIC SIGNIFICANCE TO THIS COUNTRY.

o AND THE SOLUTION OF COMPLEX PROBLEMS OF SAFETY AND HEALTH, INCLUDING THE DEVELOPMENT OF SAFETY SYSTEMS TO PREVENT NUCLEAR REACTOR ACCIDENTS, THE CREATION OF MICRO-BIOTIC ORGANISMS TO CONTAIN AND ELIMINATE HAZARDOUS WASTES, THE DEVELOPMENT OF FUTURE GENERATIONS OF CAT SCANNERS, TO ASSIST IN MEDICAL DIAGNOSIS, AND THE SOLUTION OF HUNDREDS OF BIOMEDICAL, GENETIC, MOLECULAR AND CONTAGIOUS DISEASE PROBLEMS.

IN SHORT, SCIENTIFIC SUPERCOMPUTERS HAVE BECOME TREMENDOUSLY VALUABLE IN SOLVING COMPLEX PROBLEMS VITAL TO OUR NATIONAL SECURITY, OUR ECONOMIC AND COMMERCIAL LEADERSHIP IN THE FREE WORLD AND THE HEALTH, SAFETY AND WELL-BEING OF OUR CITIZENRY. IT WOULD BE DISTRESSING, TO SAY THE LEAST, IF AMERICA LOST ITS SUPERCOMPUTER LEADERSHIP.

* * *

AT DENELCOR, WE HAVE NO INTENTION OF GIVING UP OUR LEADERSHIP POSITION TO THE JAPANESE OR ANYONE ELSE. WE MAY NOT BE AS WELL KNOWN AS SOME OTHERS IN THE INDUSTRY, BUT I AM PROUD TO REPORT TO THIS DISTINGUISHED PANEL THAT RECENTLY WE MADE WHAT MAY BE THE MOST SUBSTANTIAL "BREAK THROUGH" IN THE HISTORY OF SUPERCOMPUTER TECHNOLOGY.

When the Japanese embarked on the so-called National Super-Speed Computer Project 18 months ago, we were already four years ahead of them in the parallel processing architecture and were in the process of manufacturing our first parallel processor computing system for the U.S. Army Ballistics Research Laboratory at Aberdeen, Maryland, under the dedicated sponsorship of Dr. Eichelberger. Installed last fall, our 4-Processor "HEP", system is now recognized as the world's fastest scalar computer, with a capacity to process 40 million instructions per second.

In developing the "HEP" system, Denelcor effectively overcame the traditional problem encountered in multiple unit systems, namely the loss of a linear increase in speed and performance when each processor unit is added to the system. Our successful application of parallel process architecture has enabled us to put multiple units together and achieve a direct linear increase in speed and performance. A 16-unit HEP today makes Denelcor's system significantly faster and more powerful than any existing supercomputer of its type on the market.

RECENTLY, WE COMPLETED THE BASIC DESIGN STUDY OF DENELCOR'S HEP 2 SYSTEM, AND IN THAT PROCESS WE MADE A MONUMENTAL BREAKTHROUGH. OUR NEW SYSTEM, IN ITS LARGEST CONFIGURATIONS, WILL BE CAPABLE OF PROCESSING IN EXCESS OF 12 BILLION INSTRUCTIONS PER SECOND. THAT IS ABOUT 500 TIMES FASTER AND MORE POWERFUL THAN ANY CURRENTLY AVAILABLE SUPERCOMPUTER, AND ABOUT HALF THE PROJECTED LEVEL THAT THE JAPANESE ARE STRIVING TO REACH BY THE END OF THE DECADE. DENELCOR WILL HAVE ITS HEP 2 SYSTEM ON THE MARKET BY EARLY 1986, AND EXPECTS TO CONTINUE THE NATURAL EVOLUTION OF ITS ARCHITECTURE TO EXCEED THE STATED JAPANESE NATIONAL GOAL.

WE ACHIEVED THESE RESULTS LARGELY THROUGH OUR OWN INTERNALLY FINANCED RESEARCH AND DEVELOPMENT ACTIVITIES AND WITHOUT GOVERNMENT ASSISTANCE OR FUNDING. WE BELIEVE THERE IS TREMENDOUS POTENTIAL FOR THE USE OF OUR SUPERCOMPUTER SYSTEMS.

* * *

THIS BRINGS ME TO THE MOST IMPORTANT REMARKS I'D LIKE TO MAKE TODAY - COMMENTS CONCERNING THE ROLE GOVERNMENT CAN PLAY IN MAINTAINING AMERICAN PRE-EMINENCE IN SUPERCOMPUTER TECHNOLOGY. AS A BUSINESSMAN AND COMPUTER DEVELOPER, I AM LESS CONCERNED ABOUT THE DANGER OF LOSING OUR LEADERSHIP IN THE DESIGN OF SUPERCOMPUTERS THAN I AM ABOUT THE DANGERS OF LOSING IS OUR LEADERSHIP IN SUPERCOMPUTER USAGE.

SMALL COMPANIES LIKE DENELCOR ARE IN AN EXCELLENT POSITION TO DEVELOP IDEAS -- NEW CONCEPTS AND INNOVATIVE DESIGNS - THROUGH THE PRIVATE INVESTMENT. BUT IF OUR GOOD IDEAS ARE TO BE OF ANY LASTING VALUE, THEY MUST BE EXPLOITED FULLY, AND FULL EXPLOITATION IS GENERALLY BEYOND THE MEANS OR EVEN THE PRIMARY INTERESTS OF COMPANIES SUCH AS OURS.

DURING THE 1950'S AND 1960'S THE GOVERNMENT PLACED A HIGH PRIORITY ON THE EXPLOITATION OF SCIENTIFIC IDEAS, AND WE BELIEVE THIS COMMITMENT SHOULD BE RENEWED IN THE 1980'S. LET US LOOK BEYOND OUR PRESENT INTERESTS IN PRODUCTION COMPUTERS, ESSENTIAL AS THEY MAY BE, AND BEGIN PREPARING FOR THE FUTURE.

THE MOST EFFECTIVE WAY GOVERNMENT CAN FULLY EXPLOIT OUR PRESENT SUPERCOMPUTER PRE-EMINENCE IS TO PUT OUR MACHINES TO USE IN SOLVING THE IMMENSELY COMPLEX PROBLEMS OUR NATION AND WORLD FACE TODAY. IN SHORT, I URGE GOVERNMENT TO RENEW THE NATIONAL COMMITMENT TO RESEARCH AND DEVELOPMENT.

I URGE ALL OF YOU IN GOVERNMENT TO PROVIDE THE LEADERSHIP OUR NATION NEEDS TO KEEP AMERICA IN THE FOREFRONT OF SCIENTIFIC AND TECHNOLOGICAL ADVANCEMENT, AND TO ENCOURAGE AND TRAIN NEW GENERATIONS OF SCIENTISTS AND ENGINEERS WHO CAN MAXIMIZE THE PROBLEM-SOLVING CAPABILITIES OF OUR ADVANCED SCIENTIFIC COMPUTING SYSTEMS.

SPECIFICALLY, WE AT DENELCOR URGE YOU TO:

1) TAKE THE LEADERSHIP ROLE IN TO FORMULATING A NATIONAL PLAN TO ACHIEVE SPECIFIC SCIENTIFIC AND TECHNOLOGICAL OBJECTIVES OVER THE NEXT ONE OR TWO DECADES, AND GRANT THE DEPARTMENT OF ENERGY RESPONSIBILITY FOR DIRECTING A COORDINATED, YET DISTINCTLY INDEPENDENT, GOVERNMENT, BUSINESS AND EDUCATIONAL EFFORT TO IMPLEMENT OUR NATIONAL PLAN. IF WE IN BUSINESS KNOW WHAT OUR NATIONAL PRIORITIES ARE, WE CAN ACHIEVE THEM.

2) INCREASE FUNDING FOR RESEARCH AND DEVELOPMENT PROJECTS, ENHANCED SUPERCOMPUTER CAPACITY AND STAFFING AT THE NATIONAL LABORATORIES OF THE DEPARTMENT OF ENERGY, THE RESEARCH FACILITIES OF THE DEPARTMENT OF DEFENSE, THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION, NATIONAL SCIENCE FOUNDATION AND IN THE INTELLIGENCE COMMUNITY. GOVERNMENT SUPPORT OF OUR NATION'S MOST TALENTED, CAPABLE AND SOPHISTICATED SUPERCOMPUTER SCIENTISTS IS THE MOST EXPEDIENT AND LEAST EXPENSIVE WAY TO PROVIDE SUPPORT TO COMPLEX PROBLEM SOLUTION THROUGH SUPERCOMPUTER RESEARCH.

3) OPEN YOUR DOORS AND PROVIDE SUPERCOMPUTER ACCESS TO OTHER INTERESTED SCIENTISTS AND ENGINEERS BY ESTABLISHING RESEARCH, DEVELOPMENT AND PROBLEM SOLVING NETWORKS REACHING OUT FROM THE NATIONAL LABORATORIES TO REGIONAL GOVERNMENT FACILITIES AND BEYOND TO PRIVATE RESEARCH AGENCIES AND OUR NATION'S UNIVERSITIES.

4) Encourage the development of component and packaging technology for U.S. supercomputers of the 1980's and 1990's to alleviate our present and projected dependence on foreign suppliers of computer components.

5) Stimulate development of software for high performance computing systems, including the design of algorithms, numerical analysis and computational mathematics to maximize the performance potential of new generation supercomputers, and

6) Spark the imagination of young scientists and engineers by providing opportunities for training and real-time use of supercomputers through the establishment of fellowships, loans, and special work programs between government agencies and universities.

By taking these actions today, you in government will secure American supercomputer leadership in the years ahead. We in industry can maintain America's lead in the design and manufacture of super scientific computing systems if you in government can provide the lead in putting our supercomputers to work in scientific and technological research and development. Together we can solve the complex problems of our times.

Thank you very much.

Mr. FUQUA. Thank you very much, Mr. Hill. As soon as we conclude with the next witness, then we will have questions.

My understanding is that Mr. John F. Carlson, executive vice president and chief financial officer of Cray will present the next testimony.

Mr. CARLSON. Thank you, Mr. Chairman.

I am appearing today through prior arrangement that the committee had with Mr. Rollwagen, who was supposed to be here today. I appreciate the opportunity to be here in his absence.

Mr. Chairman, members of the committee, I am the chief financial officer and executive vice president for Cray Research and today Cray Research is the leading supplier and developer of large scale scientific computer systems.

I appreciate the opportunity to speak to you today concerning the future and the current status of supercomputer development required to meet the challenging needs of energy research. We appreciate and commend the concern you have demonstrated by these hearings. We also appreciate your broader concern that the United States maintain its strong position of world leadership in supercomputer technology.

While I am happy to report that we at Cray Research are committed and optimistic about the future, there are some matters of deep concern which could place us at considerable disadvantage if they are not dealt with in an expeditious manner. I would like to discuss at this time some of those concerns:

Almost 40 percent of the systems that Cray Research has delivered so far are employed in some type of energy research. Most of these systems support energy research conducted by the Department of Energy national laboratories and similar facilities in France and the United Kingdom. Some of the other witnesses at these hearings today will detail the important contributions that supercomputers are making to these laboratories, both by the solution of otherwise unsolvable problems, and by allowing the laboratories to conduct their research at a dramatically reduced cost, by eliminating the need for many physical experiments.

The second area of energy research to which supercomputers are the key is in support of exploration for and develoment of petroleum resources. Late in 1980, Cray Research installed a machine at the Atlantic Richfield Oil Co. Since that time, seven additional oil companies have installed supercomputers. Using advanced modeling technology and seismic signaling processing techniques, the oil companies expect to demonstrate dramatic dollar savings in both exploration and reservoir management.

In addition to direct energy research there are a number of related opportunities which are highly dependent upon current and future generations of supercomputers. One of the most important lies in the area of computational fluid dynamics. We expect that this will be discussed by Dr. Kerrebrock of NASA, during his testimony.

Supercomputers have already made significant and well documented contributions to energy and general scientific research. However, we at Cray Research believe that the full realization of the potential impact that advanced computing could have on the Nation's economy and security is just beginning to be realized.

The results we have obtained using this generation of supercomputers has spawned new methods fundamental to the birthing of emerging technologies. Some of these technologies have already reached the national agenda, such as the recently announced presidential initiative to design a space based antiballistic missile system. Access to state of the art systems will be the key to advances in technology of both societal and political importance, including biotechnology and medical research.

In recognition of these facts, other Western countries are moving to capitalize on the potential that supercomputing offers. The French have already initiated a national supercomputer development project, and have begun to broadly apply supercomputers. The British are not far behind. You are already well aware of the efforts of the Japanese in this area.

Mr. Chairman, the world is on the verge of a major computing revolution, as supercomputing is applied across broad industrial and scientific bases. The only question at this juncture is which nation will choose to lead this revolution.

The cost tradeoffs in using supercomputers are dramatic. If an oil company, for example, can just be spared the costs related to drilling a single dry hole, the entire supercomputer installation will be paid for several times.

If a major aircraft or automobile company can improve the automobile design through simulation and also shorten the cycle in the development of new automobiles, the savings to the nation in domestic fuel reserves can be staggering.

In addition, we have not yet begun to reap the potential benefits of the use of supercomputers in the biotechnology, medical care delivery systems, and a vast array of operational defense problems.

As we push more into space, the importance of processing huge amounts of data in a short period of time becomes very, very important.

Supercomputers have already been used and shown to be very effective in the evaluations and capabilities not realized or realizable by any other type of computer systems. Even in the area of artificial intelligence, which is really not the subject of this discussion, but which is the subject of considerable research around the world, supercomputers can provide a very effective tool.

As I said earlier, we are really on the verge of a major worldwide computing revolution in which supercomputers stand to play a very fundamental role. Leadership in supercomputing technologies, and in the aggressive implementation of that technology, can play a significant role in the future of both our basic industries and so-called new technology frontiers.

Cray Research is the world leader in supercomputer technology, not merely from a desire to be better than anyone else, but because we are seriously addressing important problems such as those found in energy and whose solutions are of national consequence.

Our challenge today is to ensure that we can continue to keep that leadership position. In meeting the challenge, we must not focus upon the plans or accomplishments of others, our primary motivation as a company is not to defeat the Japanese, but for the nation and for Cray the challenge is to meet society's needs.

We are confident that the United States is still in the leadership position in supercomputers at this time. Although some of the announcements by Japanese manufacturers have indicated impressive capabilities on paper, the real performance tests for supercomputers is their capability to get the scientific job done. It will still be quite a while before that can be evaluated, but we expect the current and near term offerings of Cray Research will equal or surpass the Japanese challenge.

Cray Research is ready to meet the supercomputing needs of our research and industrial community. In order to best accomplish our work, however, we need an atmosphere which will continue to spark and reward the enterprise which has fostered the growth of our industry in the past 25 years.

Until recently the success of Cray Research has depended largely upon factors within our control, such as an entrepreneural management style, and commitment to research. We have worked hard to optimize these in terms of our formula for success. Today, however, there are other factors, such as national economic and trade policy, semiconductor availability, and domestic supercomputer policy, which have risen in importance.

With the serious challenges from the Japanese, the British and French, many of our efforts hinge upon technology related economic policy. Since we at Cray Research devote large financial resources to research, our efforts have benefitted substantially from recently enacted temporary research and development tax cuts. These incentives need to be strengthened and enacted permanently.

Because our financial success and our position of world leadership is dependent upon foreign sales, we are very sensitive to changes in export administration and West/West technology transfer policy. We need a consistent, well managed and reasonable policy which will protect national security without harming our national economic assets.

Finally, efforts to enhance the quality and the quantity of trained scientists and engineers will all have a substantial impact upon the future of U.S. leadership in supercomputing.

Lack of high speed semiconductors are currently the Achilles Heel of the U.S. supercomputer effort. All of our product development projects have suffered in one way or another due to the lack of high speed reliable semiconductor parts. This is particularly true concerning logic chips, which we are very reticent to continue buying from the Japanese. We already purchase a large percentage of our memory components from Japanese suppliers.

We are concerned about continued future reliance upon the Japanese semiconductor industry for chips which will be required to build our next generation of computers. If this happens, U.S. leadership would begin to depend upon the willingness of Japanese suppliers to sell us the most advanced chips while at the same time making these chips available to our Japanese competitors.

Our second concern with this situation is that if we continue to buy logic chips from the Japanese, we will be forced to transfer to our supplier the proprietary "masks" which we have developed for our products. This type of technology transfer could end up being detrimental to the U.S. competitiveness and to our national security.

As we look at future project developments at Cray Research, we are very excited. By the end of next year we will have introduced in the last four years four new products. More importantly, in 1984 we expect to deliver the first Cray-2 system. This system will usher in a number of new technologies which will be required to meet our goals for the second half of the decade.

Mr. Chairman, we are very encouraged with the current state of supercomputer development so far as it relates to bringing the fastest possible hardware to the marketplace in the shortest period of time. If the market continues to be robust and some of the software applications difficulties can be overcome, it is not overly optimistic to assume that a system which Cray Research can deliver in 1990 will perform at speeds two orders of magnitude faster than the current Cray-1 and have memory sizes which are expected to increase at the same rate.

The Department of Energy laboratory has played a very crucial partnership role in the development of supercomputers since Mr. Cray began his career. In addition to providing an immediate market for successful products, the Department has sponsored software development, applications research, and an atmosphere of close cooperation with Cray Research.

In many respects, this particular vendor/customer relationship has been a critical key to the success of Cray Research's efforts to develop another generation of supercomputers beyond our original Cray-1. I am aware that this committee has encouraged the Department to continue in this crucial role and commend you for this action.

In his testimony, I expect that Dr. Trivelpiece of the Department of Energy will point out the importance of strengthening this partnership if we are to meet the challenges at hand. We at Cray Research strongly echo this statement.

We are entering the new generation of development in a number of dimensions. Lack of parts is impeding progress and the bulk of both spending and competition is coming from abroad. The importance of supercomputing to the next generation of industrial technology has elevated supercomputing from a specialized domain of energy research and limited set of our customers to a technology fundamental to the competitiveness of U.S. industry.

Finally, although we applaud the efforts of the Department of Energy and others to assist in supercomputer development, we are concerned that this zeal could result in efforts to attack the wrong problems. Specifically, I offer the following recommendations as you consider the proper course:

First, Insure that the Department is sufficiently funded to pursue its charter. Using supercomputers, the Department is able to leverage taxpayer investment better than with any competing technology. Under no circumstances should the computer equipment budget for approved operational requirements be curtailed in favor of the development projects or other research projects.

Second, enhance the Department's capability to acquire computers strictly for software and algorithm research. The most advanced system available should always be available for general research. This should not be limited, of course, to experimental systems but should include advanced production systems such as the Cray X-MP and Cray-2.

Third, encourage and provide extra appropriations for university access under Department sponsorship. One way to develop and facilitate this access would be to move Cray-1 systems from larger laboratories to smaller laboratories, as the Cray-1's are replaced with X-MP, Cray-2's or other systems.

Fourth, encourage continued close partnership between universities, industries, and laboratories. We at Cray Research are committed to working closely with the Department to define the best mechanism for this purpose.

In conclusion, Mr. Chairman, I am pleased to appear before you this afternoon as a representative of our company and our customers. Even more important, however, I am appearing as a representative of our technology, and those who are directly responsible for the American dominance of this vital technology: Seymour Cray, chief designer of the Cray-1; Les Davis, Cray Research executive vice president for manufacturing, research and development, and the individual who has delivered the Cray machines to users; and Dr. Steve Chen, the chief designer for the first parallel processing supercomputer, Cray X-MP, which is currently the world's fastest and largest capacity supercomputer.

Only individuals like these very select individuals develop supercomputers capable of commercial and scientific success. Development of successful supercomputers is an art and discipline which appears to emanate from a critical mass of talent, singlemindedness, support, and an opportunity to win. Although we may address issues of importance to individual companies, or consumer constituencies, the success of these hearings and any other action which the Congress or the administration takes will ultimately be measured by how this action fosters an atmosphere in which these individuals and others waiting to emerge can flourish and deliver new generations of systems to the open marketplace.

If Cray Research can provide an environment where new creative talent can emerge and, at the same time, provide an environment where an exceptional individual like Seymour Cray can continue to be creative, we will have done something very special. We believe we have begun to reap the benefits of that accomplishment and we are committed to helping our Nation to do the same.

Thank you.

[The prepared statement of Mr. Carlson follows:]

PREPARED STATEMENT OF JOHN F. CARLSON, EXECUTIVE VICE PRESIDENT AND CHIEF FINANCIAL OFFICER, CRAY RESEARCH, INC.

MR. CHAIRMAN AND MEMBERS OF THIS DISTINGUISHED COMMITTEE. I am John F. Carlson, Executive Vice President and Chief Financial Officer of Cray Research, Inc., the world's leading supplier and developer of supercomputers and supercomputing technology. Since it was founded in 1972, Cray Research has grown to over 1400 employees located across the United States and in five foreign countries (United Kingdom, West Germany, France, Netherlands, and Japan). By the end of 1983, we will have added two more countries -- Canada and Sweden -- to our foreign installed base. We currently manufacture four major products, including two types of central processors, the CRAY-1M and the CRAY X-MP. The CRAY-2 is in final stages of development and will be available in late 1984. To date, we have installed fifty-two CRAY-1 supercomputers, approximately eighty percent of all current generation supercomputers worldwide. By year end, we expect to install 14 new systems including six of the new CRAY X-MP's.

Thank you for this opportunity to speak with you concerning the current and future status of the supercomputer development required to meet future challenges in energy research. We appreciate and commend the concern you have demonstrated by these hearings. We also appreciate your broader concern that the United States maintain its strong position of world leadership in supercomputing technology. While I am happy to report that we at Cray Research are excited and optimistic about the future, there are some matters of deep concern which could place us at a considerable disadvantage if they are not dealt with expeditiously. I shall discuss these concerns later in this testimony.

For the purpose of this testimony, "supercomputer" refers to very high speed general purpose computer systems such as those produced by Cray Research and used by the Department of Energy Laboratories at Los Alamos and Albuquerque, New Mexico; and at Livermore, California. We are specifically not referring to Artificial Intelligence or "Fifth Generation" systems which are the subjects of a major new research program recently proposed by the Department of Defense Advanced Research Project Agency (DARPA). Since there has been considerable confusion in the press, this distinction is important.

IMPORTANCE OF SUPERCOMPUTERS

BASIC TOOL FOR ENERGY RESEARCH

Almost forty percent of our systems are used to support some type of energy research. Most of these systems support energy research conducted at DOE National Laboratories, and similar facilities in Germany, France and the United Kingdom. Other witnesses at these hearings have already detailed the important contribution that supercomputers are making at these laboratories -- both by the solution of otherwise intractable problems and by allowing the laboratories to conduct their research at a dramatically reduced cost by eliminating the need for many physical experiments.

A second area of energy research to which supercomputers are the key is that in support of exploration for and development of petroleum resources. Beginning with Atlantic Richfield Oil Company, which acquired a CRAY-1 in late 1980, seven major oil companies have installed supercomputers. Using advanced modeling and seismic signal processing techniques, oil companies expect to demonstrate dramatic dollar savings in exploration and reservoir management. These savings translate directly to the American consumer; better reservoir management allows us to conserve our petroleum reserves.

In addition to the direct energy research, there are a number of related opportunities which are highly dependent upon current and future generations of supercomputers. One of the most important lies in the area of Computational Fluid Dynamics, discussed earlier by Dr. Kerrebrock, of NASA. In addition to the significant impact upon the entire process of engineering design, recent results from the use of the CRAY-1 supercomputer by the aerospace community have demonstrated the potential for designing air frames with significant savings in total fuel consumption, translating into direct savings to both the public and private sector and preservation of national fuel reserves. We expect the impact of these savings to grow as supercomputers are introduced into the automobile and shipbuilding industries.

FUNDAMENTAL TOOL FOR DEVELOPMENT OF NEW TECHNOLOGIES

Supercomputers have made significant and well-documented contributions to energy and general scientific research. However, we have only begun to realize the potential impact that advanced computing could have upon our Nation's economy and security. The results we have obtained using this generation of computers have spawned new methods fundamental to the birthing of emerging technologies. Some of these technologies have already reached the National agenda, such as the recently announced Presidential initiative to design a space-based anti-ballistic missile system. Access to state-of-the-art systems will be key to advances in technologies of societal and political importance including biotechnology and medical research. Recognizing these facts, other Western countries are moving to capitalize upon the potential that supercomputing offers. The French have instituted a national supercomputer development project and begun to broadly apply supercomputers. The British are not far behind, both in developing their understanding of supercomputing and in preparing for national initiatives. This committee is well aware of Japanese efforts.

CONTINUED WORLD LEADERSHIP A PRIORITY

Mr. Chairman, the world is on the verge of a major computing revolution as supercomputing is applied across a broad industrial and scientific base. The only question at this juncture is which nation will choose to lead this revolution.

The cost trade-offs are dramatic. If an oil company can be spared drilling a single dry hole, the entire supercomputer installation will be paid for several times over. If a major aircraft or a new automobile model can be made as little as 2 or 3 percent more fuel-efficient, the savings to the nation in domestic fuel reserves can be staggering. We have not yet begun to reap the benefits of the potential use of supercomputers in biotechnology, medical health care delivery systems, socio-economic modeling and a vast array of operational defense problems. As we push more into space, the processing of huge amounts of data in a short time becomes even more important. Supercomputers have already been shown in some evaluations to have considerable efficiencies and capabilities not realizable in any other class of system. Even in the area of artificial intelligence, which is a subject of considerable research around the world, supercomputers stand to be a powerful tool. Artificial intelligence algorithms are known to be very complex, but the payoff in certain areas could be substantial if these algorithms can be developed in a reasonable amount of time.

As we approach the environment of the year 2000, we should not be surprised that Japan, France and the United Kingdom have all recognized the importance of supercomputing to the extent that they have initiated national programs of varying scopes to deal with the issue. Many have questioned why the Japanese would enter supercomputing since their normal strategy is to target industries with broad commercial significance. While it is possible that the venture into supercomputing has been more for prestige, I believe that history will record the venture as a very shrewd one. When Fujitsu and Hitachi announced their supercomputers last year they projected deliveries to approximately 30 domestic customers in the first 18 months, or more new customers than Cray Research expects worldwide during the same period. At this time there are only two current generation supercomputers in Japan. If these Japanese manufacturers are successful, we can expect to see a computing revolution throughout Japanese industry that could further threaten U.S. industrial competitiveness.

As these facts indicate, the importance of leadership in supercomputing technology goes far beyond the already well-established links to Defense and National Security. As I said earlier, we are on the verge of a major worldwide computing revolution in which supercomputers stand to play a fundamental role. Leadership in supercomputing technologies, and in the aggressive implementation of that technology, could play a significant role in the future of both our basic industries, and new high technology frontiers.

Cray Research is the world leader in supercomputer technology, not merely from a desire to be better than anyone else, but because we are seriously addressing important problems, such as those found in energy whose solutions are of National consequence. We are addressing these problems in a constructive, imaginative and American way -- and the result has been excellence. The United States as a Nation and Cray Research as a private enterprise have each demonstrated their ability to achieve world leadership. But leadership does not guarantee success. True success is sustained achievement. Our challenge today is to ensure that we can continue to keep that leadership. In meeting that challenge, we must not focus upon the plans or accomplishments of others. Our primary motivation is not to defeat the Japanese. For the Nation and for Cray the current challenge is to meet society's needs. In order to adequately address the future needs of our industries, our technology, and the National Security, we must foster an environment of dynamism in research and development. We must define the problems to be solved and focus upon the solutions rather than the competitors.

We are confident that the United States is still the leader in supercomputing at this time. Although some of the announcements by Japanese manufacturers have indicated impressive capabilities on paper, the real performance test for supercomputers is their capability to get the scientific job done. It will still be quite a while before that can even be evaluated, but we expect that the current and near-term future offerings of Cray Research will equal or surpass the Japanese challenge.

Cray Research, and, we believe the Nation, is ready to meet the supercomputing needs of our research and industrial community. In order to best accomplish our work however, we need an atmosphere which will continue to spark and reward the enterprise which has fostered the growth of our industry for the past twenty-five years. The formula for success is far from simple. Until recently, the success of Cray Research has depended largely upon factors within our control, such as management style, internal commitment to research, etc. We have worked hard to optimize those terms of the formula. We have promoted innovative management. We have committed 15-17% of our gross annual revenues to research and development, yielding a total supercomputer development budget comparable to the entire published budget for the Japanese Supercomputer Development Project. We have created an environment where other supercomputer designers can flourish side by side with Seymour Cray. We have marketed aggressively to foreign customers located within Allied nations to the point that over forty percent of our new revenues are coming from foreign sales. We have sought to establish and maintain close contact with our customer base to ensure that we are sensitive to their future technology needs. These efforts have proven fruitful thus far, and we believe we have secured supercomputing preeminence for the next generation of computers.

Today, however, other terms of the formula -- national economic and trade policy, semiconductor availability and domestic supercomputer policy -- have risen in importance. In addition, the specter of successful industry targeting by the Japanese, French, and British has dramatically increased the potential negative impact of any artificial barriers imposed by Federal policy.

NEED FOR TECHNOLOGY-ORIENTED ECONOMIC AND TRADE POLICY

With the serious challenges from the Japanese, British, and French, many of our efforts hinge upon technology-related economic policy. Since we devote large financial resources to research, our efforts have benefited substantially from recently enacted temporary research and development tax incentives. These incentives need to be strengthened and enacted permanently. Because our financial success and our position of world leadership is dependent upon foreign sales, we are very sensitive to changes in export administration and West-West technology transfer policy. We need a consistent, well-managed, and reasonable policy which will protect National security without harming our National economic assets. Finally, efforts to enhance the quality and quantity of trained scientists and engineers will all have a substantial impact upon the future of U.S. leadership in supercomputing. These issues have been the subject of other hearings, and I shall not address them further today.

NEED FOR ADVANCED SEMICONDUCTORS

Lack of high speed semiconductors is currently the Achilles heel of the U.S. supercomputer effort. Product development projects have been delayed or even suspended due to the lack of new high-speed reliable semiconductor parts. Even the major government-sponsored semiconductor research programs - the VHSIC program and the DARPA Galium Arsenide program - cannot be expected to yield any major advantages to the U.S. supercomputer industry.

Cray Research already purchases most of its memory components from abroad. We are concerned that we will have to rely more and more upon the Japanese semiconductor industry for the high speed components required to build our computers. If this happens, U.S. leadership would begin to depend upon the willingness of Japanese suppliers to sell us the most advanced chips at the same time that these chips are available to our Japanese competitors.

The problem is more serious in the area of logic components. If we buy our logic components from the Japanese we will be forced to transfer to our supplier the proprietary mask design which reveals exactly how we will be using the chip. This type of technology transfer could be detrimental to both U.S. competitiveness and National Security, if it is forced upon us by lack of domestic supply.

Having addressed what we consider the primary impediments to U.S. leadership in supercomputing, I will now turn to the final area of the "formula for success" — that of the Federal role in fostering supercomputer development.

THE FEDERAL ROLE IN FOSTERING SUPERCOMPUTER DEVELOPMENT

CURRENT STATUS OF SUPERCOMPUTER DEVELOPMENT

We are currently in an exciting phase at Cray Research. By the end of next year, we will have delivered a major new product each year for four years. In 1981, we delivered the first Input/Output subsystem, a special high speed processor which dramatically enhanced the capability of the CRAY-1 supercomputer. In 1982 we delivered the first CRAY X-MP parallel processing supercomputer to our internal data center, marking the first time in over twenty years that someone other than Seymour Cray had designed a commercially successful supercomputer faster than Cray's latest system. We also delivered the first Solid State Storage Device, a high speed, 256 million byte, auxiliary memory which connects to the CRAY X-MP via a communications link capable of operating at speeds up to ten gigabits per second. In 1983, we will deliver the first CRAY 1/M, a follow-on to the CRAY-1 processor which uses MOS (instead of ECL) memory, thus reducing the system price as much as fifty percent. In 1984, we expect to deliver the first CRAY-2 system, which will usher in a number of new technologies required to meet our goals for the second half of the decade.

We are very encouraged with the current state of supercomputer development. If the market continues to be robust, and if some of the software and applications difficulties can be overcome, it is not overly optimistic to assume that the system which Cray Research delivers in 1990 will perform at speeds at least two orders of magnitude faster than the CRAY-1; memory sizes are expected to increase at the same rate.

In the applications area, we are not quite as encouraged. The record of the last five years indicates that other than in environments such as the Department of Energy Laboratories, the U.S. research and engineering community has been slow to implement supercomputing technology. Even the Defense and aerospace communities, which have in the past aggressively applied new computers, have lagged seriously behind their counterparts overseas. Since the success of new computing technologies is heavily dependent upon their application in a user environment, lack of aggressive use of new generations of supercomputers could impede attempts to speed up the development cycle.

In the software area, there is a need for parallel processing breakthroughs in software technology. There is a critical need to move the state of the art in this area as quickly as possible. Now that the CRAY X-MP is available, it is possible for researchers to simultaneously investigate parallel processing and high speed vector processing. Unfortunately, unless action is taken in the near future, this unique tool, like supercomputers before it, will be largely inaccessible to the vast University research community.

The greatest immediate challenge from the Japanese as well as from Western Europe is their aggressive implementation of new technology. In the United States, supercomputers are largely inaccessible to the vast majority of defense, industrial and academic scientists and engineers. The picture is quite different in Europe and Japan, where the governments have long recognized the importance of supercomputers to their world competitive positions.

THE PROPER ROLE OF THE DEPARTMENT OF ENERGY IN SUPERCOMPUTER DEVELOPMENT

The Department of Energy Laboratories have played a crucial partnership role in the development of supercomputers since Mr. Cray began his career. In addition to providing an immediate market for successful products, the Department has sponsored software development and application research in an atmosphere of close cooperation with Cray Research. In many respects, this particular vendor-customer relationship has been a crucial key to the success of Cray Research's efforts in developing another generation of supercomputer beyond the original CRAY-1. I am aware that this Committee has encouraged the Department to continue in this crucial role and commend you for this action.

In his earlier testimony, Dr. Trivelpiece of the Department of Energy pointed out the importance of strengthening this partnership if we are to meet the challenges. We at Cray Research strongly echo this statement. We are entering a new generation of development in a number of dimensions. Ten years ago, when Mr. Cray was building the first CRAY-1, he was able to use supporting technologies which were immediately available. There was no foreign competition, and there was no indication that future development would be limited by availability of domestic parts. Furthermore, from a National Interest standpoint, four major U.S. manufacturers were expending major resources to compete (Burroughs, IBM, CDC, and Texas Instruments). Today, the picture is quite different. Lack of parts is impeding progress and the bulk of both spending and competition is coming from abroad. In addition, the importance of supercomputing to the next generation of industrial technology has elevated supercomputing from a specialized domain of energy research and a limited set of other customers to a technology fundamental to the competitiveness of U.S. industry.

Finally, although we applaud the efforts of the Department and others to assist in supercomputer development, we are concerned that this zeal could result in efforts to attack the wrong problems. Specifically, I offer the following recommendations as you consider the proper course:

- Ensure that the Department is sufficiently funded to pursue its charter. Using supercomputers, the Department is able to leverage the taxpayer investment better than with any competing technology. Under no circumstance should computer equipment budgets designed to meet program requirements be curtailed in favor of non-program related in-house development projects, or other research. If the rate of development is to be increased, the potential market and aggressive application development must be encouraged first.

- Enhance the Department's capability to acquire computers strictly for software and algorithm research. The most advanced system available should always be available for general research. This should not be limited, of course, to "experimental" systems, but should include the newest production systems such as the CRAY X-MP or CRAY-2.

- Encourage and provide extra appropriations for University access under Department sponsorship. One way to facilitate this access would be to move CRAY-1 systems from the larger laboratories to smaller laboratories as the CRAY-1's are replaced by CRAY X-MP's, CRAY-2's or other systems.

- Encourage continued close partnership between Universities, industry and the laboratories. We are committed to work closely with the Department to define the best mechanism for this purpose.

CONCLUSION

In conclusion, Mr. Chairman, I am pleased to appear before you this afternoon as a representative of our company and our customers. Even more important, however, I am appearing as the representative of our technology and those who are most directly responsible for American dominance of this vital technology: Seymour Cray, the chief designer of the CRAY-1 and its predecessors, the Control Data 1604, 6600, and 7600; Les Davis, Cray Research Executive Vice President for Manufacturing, Research and Development, and the individual who has delivered the Cray machines to the users; and Dr. Steve Chen, the chief designer for the first parallel processing supercomputer, the CRAY X-MP, which is currently the world's fastest and largest capacity supercomputer. Individuals like these -- very select individuals -- develop supercomputers capable of commercial and scientific success. Development of successful supercomputers is an art and a discipline which appears to emanate from a "critical mass" of talent, single-mindedness, support, and the opportunity to win.

Although we may address issues of importance to individual companies or customer constituencies, the success of these hearings and any action which the Congress or Administration takes will ultimately be measured by how that action fosters an atmosphere in which these individuals, and others waiting to emerge can flourish and deliver new generations of systems to the open marketplace. If Cray Research can provide an environment where new creative talent can emerge, and at the same time provide an environment where an exceptional individual like Seymour Cray can continue to be creative, we will indeed have done something very special. We believe we have begun to reap the benefits of that accomplishment and we are committed to helping our Nation to do the same.

Mr. Fuqua. Thank you very much.

The committee is going to have to take a short recess. We have a vote going on on the floor. We will return and we will have some questions for you. I apologize for the interruption, but we have to keep a few balls juggling in the air at the same time.

[Recess.]

Ms. Lloyd [presiding]. The subcommittee will come to order.

In the interest of time I will move ahead with the question-and-answer part of this phase of our hearing while we are awaiting our chairman to come back. We appreciate your very fine testmony and the contribution that you have made to our hearings. Of course, one of the objectives of the hearing is to establish how we can maintain our leadership in the state-of-the-art computer field, but also I believe you brought out, Mr. Carlson, that we sustain our position of leadership as well.

As I move the thought in this direction I think we should look at this in a very objective manner in seeing in what areas we can enhance our position, not only in leading but sustaining our leadership role. I think we should look at the Lax report of last December that stated that the likelihood of maintaining this leadership was very slim. Would you agree with that statement, Mr. Sumney?

Mr. Sumney. Yes, I think that is reasonably accurate.

Ms. Lloyd. Maintain the leadership, you feel that we can digress from a position of leadership, and if so what would you do to enhance our position or to stabilize where we are at the present time? Would it be in the field of education, enhancing the use of computers, or antitrust laws? What would you see as one of the major objectives?

Mr. Sumney. I think that the SRC provides sort of an example of something that seems to work reasonably well. What we have shown so far is that the industry is willing to engage in a new mode of activity, that of sharing the costs of generic research. Our present approach is to saturate the university community first because of the need for better and more highly trained graduates.

Ms. Lloyd. Back up just a minute. How much of a factor is generic research in moving ahead with our computer technology?

Mr. Sumney. I think that the gentleman that testified for Cray indicated that high-speed semiconductors and integrated circuits were a key part of superior computer development. Almost everything we do ties into higher densities or high-speed, leaving the fundamental research necessary to do the development for the kinds of things that superior computers are going to need. The R&D is becoming expensive, and the cooperative route appears to be working well. The missing link is the Government role. We are working with universities and industry and are exploring in a tentative way methods of involving the Government in some of the things that we are doing to form sort of a three-pronged attack on the problem, and as we do that we will move from basic and applied research into development.

Ms. Lloyd. Would you certainly look with favor upon DOE being more involved in the research activities?

Mr. Sumney. Yes.

Ms. Lloyd. In what way, funding or actually working for policy?

Mr. **Sumney**. Peripherally we have worked with the national laboratories. We have talked to Albuquerque and we are very familiar with the Sandia laboratory. I think that the availability of those facilities in some way can be very useful to what we are trying to do. Sandia has established a premier integrated circuit which is valuable. So from that limited standpoint, from a research resource it seems to the laboratories could be something that we could do and tie that capability into the research that we are doing in the university.

Ms. **Lloyd**. Thank you very much.

Mr. Carlson, you touched upon some of our antitrust laws I believe in your testimony. This subcommittee has had many articles recently come to their attention that cite the problems posed by antitrust laws encouraging innovations in the electronics industry. Do you believe that we should review the effects of our antitrust laws in encouraging R&D in this effort? We would like your opinion on that.

Mr. **Carlson**. I believe the antitrust laws should be reviewed. In addition, we would like to see fundamental support of research and development, with the research and development tax credit as being one way to handle that.

Ms. **Lloyd**. Mr. Sumney and Mr. Hill, would you care to comment on our antitrust laws? Do they actually discourage innovation in the electronics industry?

Mr. **Hill**. I think recent indications are a little bit of a lessening with the MCC formulation and shared research, which is very European style, I think is excellent. Companies like Denelcor are small and we do not have a large resource pool. If there were more trade associations that we could interact with such as SRC and MCC without having to pay the high entry fee, it is a rich company's game so far, could be beneficial to us and there could be focused research that we could take advantage of. High-speed circuits, clearly a concern. Computer architecture evolves on the progress made in logic and memory speed. I support Mr. Carlson's comments.

Ms. **Lloyd**. What would you suggest might be helpful in designing a comprehensive national policy to really move toward successful technology development in microelectronics or superior computers? Trade policy? Mr. Hill?

Mr. **Hill**. Trade policy?

Ms. **Lloyd**. Design a national policy?

Mr. **Hill**. I think one of the major elements that is missing is the leadership that is required, and I propose the Department of Energy, because in my opinion they are the preeminent computer users and architects in the Government circle. A goal—you cannot lead until you know what leading is, and unless there is a goal established that says this type of platform is required by 1990 to run at these speeds or to solve these types of user problems, we in the industrial base do not have a clearcut target. I think that somehow the interagency activity going on today can result in some leadership statements. I do not believe nor would I ask the Government to support a massive infusion of Government dollars and Government control and direction into this. I believe that there can be some leadership and some stimulus, and that the industries driven by their motivation to make money and to grow can fill this void nicely.

Ms. LLOYD. You are talking more about a national policy and goals from the Department of Energy and less focus on financial incentives as well.

Mr. HILL. Yes.

Ms. LLOYD. Following on this same thought, there has been a great deal of discussion about the need to provide some form of an immunity from the sanctions of antitrust laws to electronic companies who wish to conduct R&D on a joint basis. Do you think that a proposal such as this would encourage innovative research in the electronics area?

Mr. SUMNEY. I do. We certainly established the SRC along the lines of the current antitrust regulations. We have yet to get a formal ruling that we are OK, although informally the opinion seems to be that we are. Some of the things that we have limited ourselves to because of antitrust are basic research, having the SRC open to anybody, and that means worldwide, and making the results available to everybody. Those are some restrictions that often are not in the best interests of the member companies or in the best interests of this country, and we have had to do that to stay clear of antitrust regulations. The MCC, which is a similar kind of activity, except that they are principally involved in the development of computer technology, we are principally involved in research in semiconductors as a rough cut. They have not stuck as closely to the research route as we have and perhaps are more vulnerable to antitrust problems than we are. I think that a review of the antitrust regulations and a restructuring to allow cooperative activities to take place would be to the great benefit of the country.

Ms. LLOYD. Do any of you other gentlemen like to comment?

Mr. HILL. I do not feel as strong on that issue. I believe there are some good points to some of the antitrust legislation that is in place that cause markets and competition to flourish.

Ms. LLOYD. Do you think that our policies place us at a disadvantage competitively?

Mr. HILL. No.

Mr. CARLSON. I do not think we are at a disadvantage. We are the leading country in terms of superior computers and it is best left to the individual companies through the system that is presently in place in the United States to foster that. I think that support from the Government standpoint should come through buying computers.

Ms. LLOYD. You stated that certainly we are leading in computer technology. Then you would disagree with the statement of the Lax report of last December that we are fast losing our leadership?

Mr. CARLSON. Yes, I would disagree with that comment. I think that at Cray Research we have been able to demonstrate that through the efforts of Seymour Cray and as I indicated through Dr. Steve Chan that we have been able to continue our lead in super computers, furthermore, we are committed as a company to bringing forth, as quickly as we can, the fastest technology from a super computer standpoint. We are committed as a company to spending 15 percent of all of our revenue dollars on research and development, and it is through that mechanism that we will continue to be the leader. The best role that I think the Government can play is to buy our machines to give us the revenue dollars that we can then put back into research and development.

Ms. LLOYD. One other area. Of course we are interested in the training aspect as well, so we will talk about that just a minute. I think some of the testimony stated that lack of training was a problem, and then it has been stated there is a lack of access to computers. This looks like a chicken-and-egg controversy to me. Would any of you like to comment on this?

Mr. CARLSON. I think there is lack of access. An effective way that the Government could participate is by making super computers available at the university level. I think that in the United States we lag from what is happening in Europe and Japan from this perspective. The university marketplace does not have access to super computers in the same dimension as they have in Western Europe and in Japan.

Ms. LLOYD. You feel that we should put them on the university campuses?

Mr. CARLSON. Absolutely. The earlier you can introduce the college student to the capabilities of superior computers, the better. I think that this will be a giant step toward continuing to preserve our leadership position.

Ms. LLOYD. So you would be in favor of the Government enhancing purchasing of computers?

Mr. CARLSON. Yes. The best way for the Government to support supercomputing is to buy more of what we are producing, because we are producing the fastest, with the Cray-2 and the Cray-3, which is beyond that, provide the leadership for this country through this decade.

Mr. HILL. I would like to comment on that.

Ms. LLOYD. Mr. Hill.

Mr. HILL. I think that we have to be especially aware of the fact that the university environment has to be trained how to manage supercomputer centers, that in the future they can, but I think the national laboratories have to work in conjunction with that effort to move superior computers into the national laboratory, into the university area, and that ultimately the infrastructure that we are looking for, that must be developed with this training, the next round of Seymour Crays and Burton Smiths that we have at our disposal, we do not know where they are yet. One of the things the Japanese will have because they are making access to computers, the French are, they are beginning to develop the infrastructure that we have not begun to develop. You can count the top computer scientists in this country on both hands. Access to supercomputers I support fully for the reason of developing the infrastructure that allows us to develop the next machines in the nineties and beyond.

Ms. LLOYD. Do we have enough research in our universities that can be done on a supercomputer?

Mr. HILL. I would think so, yes. A tremendous amount of softwear research must go along with the next evolution of logic machines. In parallel architecture, the development of new problem-solving algorithms, learning how to use the tools, the university community has top people that can definitely help Government and industry in developing these.

Ms. LLOYD. What size of anticipated market would be necessary to motivate manufacturers to spend additional money on supercomputers, numerical processing, or symbolic processing?

Mr. HILL. Over today's marketplace?

Ms. LLOYD. That is correct.

Mr. HILL. The frustration that I see is that there is a need today, a stated need by the users in the national laboratories to get more supercomputers to use, so I do not see an artificial need having to be created. I would say there is some large number of demand out there that the laboratories probably will be presenting in their testimony that would be focused on it. Our interest is—you have our attention today and we think that the leverage of these hearings and what we understand is going on, there will be an adequate demand for the machines.

Mr. CARLSON. I agree with that. Market size is not necessarily a factor, but that cannot be restricted to just domestic. Trade policy is an important factor, and Cray Research could not survive as a company if we were limited to domestic sales opportunities only.

Ms. LLOYD. Thank you very much.

Mr. Fuqua.

Mr. FUQUA [presiding]. What role should the Government have to play in the industry relationship as far as the state-of-the-art computer development similar to what the Japanese are doing now?

Mr. HILL. I think the role of leadership as I stated earlier, I think what is missing is the Government is an amorphous being and it does not have structure from our view or it is difficult to find the structure, and therefore the national goal that says this level of computing by this year is important to our commercial and security factors in the country—it is difficult for all the manufacturers, I am sure to find out, what exactly are the important plateaus to shoot for so we can spend our money in the right directions. That does not necessarily mean a massive infusion of money behind that, but it does mean leadership and coordination between the agencies and talking to the suppliers.

Mr. FUQUA. Do any of the other two wish to comment?

Mr. CARLSON. I think the role should continue to be one of a partnership. In terms of setting performance criteria for supercomputers, I think the competitive environment that exists in the world today will provide that. The role of the Government should be one of supporting development efforts by companies like Cray Research through the acquisition of our computers, because we are confident that we will be in a leadership position. A second Government role is that of supporting the acquisition of supercomputers by the university community, putting them on the college campuses so that students will have access to computers early in their careers. Right now that is presently not being done in the United States but it is being done in Japan, France, Germany, and the United Kingdom. If we are to stay competitive, we will set the performance guidelines and goals that we need, at least at Cray Research. The challenge we need to provide the leadership will follow as long as the systems are required.

Mr. **Fuqua**. Mr. Hill, you mentioned in your statement about—specifically some items that we had, it was item No. 2, toward the end of your statement, you were talking about increasing, enhancing supercomputer capability at the national laboratories of DOE, NSF, you mentioned NASA, et cetera, but you did not mention anything about universities. Should we—should they have access to these in these national laboratories and facilities, or should that be combined with some action on location at the universities where they can be jointly used by others through networking and other methodologies?

Mr. **Hill**. I think the program should primarily emanate from the national laboratories. The national labs have proven to be good managers of supercomputers. They know how to structure, manage, maintain them, they know how to surround them with the right people with application code writers. I think it should evolve to more advanced campuses, but I think that has to be carefully done. During the first-stage access, dialup and an ARPANET approach could facilitate that, but over a period of 2 or 3 years, not a long time, the universities could learn how to manage supercomputers and make that happen. But I think a cold water treatment of a supercomputer onto a university campus is a difficult thing to absorb without them knowing exactly what they are getting in for. Part of their problem becomes the staff and the expense support and how to run it.

Mr. **Fuqua**. Mr. Lewis.

Mr. **Lewis**. Thank you, Mr. Chairman.

I am interested in the markets for the supercomputer and what kind of markets would be available, and what the Federal Government could do as far as working or what kind of atmosphere should they apply to a supercomputer market, and how these could be used as far as R&D in all phases as well as cooperatively with students within the schools.

Mr. Hill, could you——

Mr. **Hill**. Yes. The laboratories in the last several years have been primarily using their supercomputers in a production mode, and one of our observations is that a production mode means they have got their money funneled from other agencies, they are working on a problem they know about, they have no time for experimental work, no time to develop new methodologies in problem-solving techniques that move them to the next generation of problem-solving capability.

I think that the markets are broad, petroleum, reservoir modeling, oceanography, different types of atmospheric weather forecasting, new types of CAT scanners, electronic circuit design, CAD/CAM, all types of large simulations in modeling, supersonic aircraft development—there just is not a front moving field in the industry today that does not require and need supercomputers. The lack of the infrastructure, the people that understand how to use them, the availability of the funds to put this tool in the hands of the researchers that can develop it to a usable tool so that industry can take it is the frustrating part. So there are many, many markets that can benefit by this.

Mr. LEWIS. What role do you see the Federal Government playing as far as setting the environment for the supercomputer applications?

Mr. HILL. I agree with Mr. Carlson. I think the principal mission other than providing some type of leadership in terms of what our goals should be in terms of the most powerful machine, we may not be able to achieve that, but if the Japanese achieve their goal and we fall short we could be in trouble. I think that the purchase of machines through the demand that is already there for experimentation in the areas I mentioned would be the action I would recommend. I think it would allow the commercial entities represented here to move forward, be healthy and react to the demand and reinvest our own money in research.

Mr. LEWIS. How about centralizing computers for the use of the many colleges and universities? Mr. Carlson, would you care to comment on that?

Mr. CARLSON. I encouraged making resources available to colleges and universities so that they can acquire supercomputers and have access to them at the college and university level. I think that that is a proper role for the Government to play.

Mr. LEWIS. Do you feel that supercomputers centralized in a university, in particular States, areas, or regions could be centralized for use by both Government application and private enterprise application?

Mr. CARLSON. Yes, I think that could be done.

Mr. LEWIS. What would be required to initiate and operate such type of facilities, in your judgment?

Mr. CARLSON. Well, I think that through sponsorship by the Department of Energy, a plan could be developed that would put supercomputers at some of the major colleges and universities in the country. I agree that the program could be directed by the Department of Energy to facilitate university installations that and to allow other colleges with major science and technology programs, that do not have supercomputers located on site, to access them through high speed networks.

Mr. LEWIS. I missed the comments on any antitrust implications by the various groups getting together, industry getting together or the private sector as far as applying supercomputers for time-sharing basis and things like that. Do you have any comments on that?

Mr. SUMNEY. I guess my perspective on antitrust is a little bit different than expressed by my two colleagues here. The general feeling in the semiconductor industry is that the R&D necessary to keep that industry moving ahead is going to be so expensive, a cooperative activity in research and perhaps even in development of certain products is going to be absolutely essential for the survival of that industry. That ties back clearly to the supercomputer problem that we are discussing today, because the need for higher-speed integrated circuits and semiconductors is crucial to the progress in this arena as well. So speaking with respect to antitrust, the semiconductor industry feels that a review of the antitrust regulation toward the end of facilitating cooperative R&D ventures would be very helpful to the industry and would enable it to move forward at a much quicker pace and be competitive in the world market to a much stronger degree than it is now.

If you look at some of the memory products that the Japanese, particularly in the MOS arena, specifically in the memory products, starting with the 64K-bit RAM, now moving into the 256K-bit RAM, they have moved ahead very rapidly. If we lose that market, the semiconductor industry in this country is going to be in serious trouble, and cooperative R&D appears to them to be one of the things they feel is essential to maintain their survival. That is not easy to do under the current antitrust regulations.

Mr. LEWIS. Mr. Carlson, Mr. Hill, would you care to comment?

Mr. HILL. I think I made my position clear earlier. I think the antitrust laws are adequate as they are. I think that things like MCC and SRC are probably as liberal as I would like to see them facilitated, because I think there are a lot of good aspects to the antitrust legislation.

Mr. LEWIS. Thank you, gentlemen.

Thank you, Mr. Chairman.

Mr. FUQUA. Thank you, Mr. Lewis.

Let me thank all you gentlemen for being here and for your testimony. Thank you very much.

Mr. FUQUA. Our next panel of witnesses will be Dr. Robert Cooper, Director, Defense Advanced Research Projects Agency; Dr. Alvin Trivelpiece, Director, Office of Energy Research at the Department of Energy; and Dr. Jack Kerrebrock, Associated Administrator, Office of Aeronautics and Space Technology, National Aeronautics and Space Administration.

Dr. Cooper, now that we have got you labeled right, you may proceed and we will do as we did in the first panel, we will hear your testimony and then have questions.

If you have statements that you wish to make part of the record, we will, and if you want to summarize, we would be pleased to do that also.

[The biographical sketch of Dr. Cooper follows:]

BIOGRAPHICAL INFORMATION OF ROBERT S. COOPER, DIRECTOR, DARPA

Education: Bachelor Degree, Electrical Engineering, University of Iowa, Master Degree, Electrical Engineering, Ohio State University; Doctor of Science Degree, Electrical Engineering, Massachusetts Institute of Technology.

Academic honors: University of Iowa Graduate with Distinction; Ohio State University Westinghouse Fellow; M.I.T. Ford Foundation Post Doctoral Fellow.

Experience: August 1981 to present.

Director of the Defense Advanced Research Projects Agency. Chief executive officer of an independent Defense agency responsible for research and exploratory developmental concepts that are not yet identifiable with a specific military requirement or Service mission, or which relate to the function of more than one Military Servive. Provides technical direction and establishes management policy for an agency of about 145 people conducting R&D projects totalling more than $700 million per year.

1979–81. Vice President, Engineering, Satellite Business Systems. In charge of all research and development of an advanced, digital, burst communication system employing satellites and customer premise earth stations to provide integrated voice, high speed data, electronic mail and video teleconferencing services to SBS customers.

1975–79. Director of NASA Goddard Space Flight Center. Chief executive officer of an organization employing 3,700 scientists, engineers and support personnel engaged in space related research and development projects at an expenditure rate of about $800M per year. Job involved setting policy and determining direction of major research programs, garnering NASA, White House and Congressional support for these programs and directing their execution. Responsible for launch vehicle operation employing the Thor-Delta launch vehicle at an average rate of 10–12 per year from both Eastern Test Range and Western Test Range. Also responsible for worldwide satellite tracking system and NASA communications (NASCOM). Some representative Goddard research programs were: Space Science research in astronomy, astrophysics, magnetospheric physics, infrared and ultraviolet observational equipment, high energy physics and astronomy. Applied Science research in atmospheric observation from space and weather prediction, geophysics, oceanography, land observation from space and advanced earth observational instrument development. Engineering research in spacecraft stabilization and control, propulsion, materials, data processing and analysis, and communications.

1972–75. Assistant Director of Defense Research and Engineering, Department of Defense. Set policy, established DDR&E program direction, budgets and monitored programs in Defense Satellite Systems, High Energy Laser Weapon programs, Defense Mapping Agency, and provided DDR&E coordination with all intelligence-related R&D.

1968–72. Staff member, Group Leader and Division Director at MIT Lincoln Laboratory. Participated in and managed research and development programs ranging over a spectrum of the basic and applied engineering sciences. Some example research activities were: radar discrimination of ICBM reentry vehicles, plasma physics and radar scattering from ICBM ionized wakes, infrared and optical ICBM signatures, infrared observational equipment development, high energy laser research, laser radar development and laser spectroscopy.

1963–68. Assistant Professor of Electrical Engineering, MIT and Staff Member in the Research Laboratory for Electronics. Conducted research in plasma physics and electromagnetic scattering from plasmas. Led a small research group in the study of collective plasma phenomena.

STATEMENTS OF DR. ROBERT S. COOPER, DIRECTOR, DEFENSE ADVANCED RESEARCH PROJECTS AGENCY; ALVIN W. TRIVELPIECE, DIRECTOR, OFFICE OF ENERGY RESEARCH, DEPARTMENT OF ENERGY; JACK L. KERREBROCK, ASSOCIATION ADMINISTRATOR, OFFICE OF AERONAUTICS AND SPACE TECHNOLOGY, NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Dr. COOPER. Thank you, Mr. Chairman, distinguished members. I am pleased to be here today at your invitation to discuss my view of the state of computer science research in the United States and to review broadly DARPA's plans for developing super-intelligent computers for defense.

As you know, DARPA has been a major supporter of computer science research since the early 1960's, shortly after the agency was created by President Eisenhower in the wake of sputnik. For 25 years we have played a central role in bringing new technological concepts to fruition for the benefit of defense and also, I might add, creating spinoff benefits for the civilian sector.

DARPA is known for pioneering the development of computer time-sharing in the early 1960's, computer graphics in the mid-1960's, and packet switching technology, and the ARPANET in particular, in the late 1960's and early 1970's.

During the 1970's, we led the development of network protocols, demonstrated the application of packet switching to radio and satellite systems and recently have embarked on an ambitious program to encourage more widespread use of VLSI design and fabrication by our leading universities and research labs; relying on commercial vendors for actual fabrication of integrated circuit chips.

Since our earliest days as an agency, DARPA—then ARPA—has provided a major portion of the Nation's support for artificial intelligence research. This work is aimed at developing machines that will have human-like capabilities for intelligent behavior.

As you know, this technology is expected to form the basis of the so-called fifth generation computer. We believe this field holds special promise to produce a class of very intelligent computers such as Defense will require within the next few years.

In the fiscal year 1984 President's budget submission to the Congress, we included a new program entitled strategic computing and survivability, which directly addresses this need.

Over the next few years, we plan to develop a class of super intelligent machines and to apply them to a wide range of critical defense problems.

Three initial areas have been identified as a specific focus for this work; however, we anticipate that others may also prove to be of high priority as the technology base evolves.

In particular, a Defense Science Board task force chaired by Prof. Joshua Lederberg of Rockefeller University has been established to study the area of military priorities for applications of super intelligent computers. We expect to have their results early next year.

This program represents somewhat of a departure for DARPA in that up to now funding limitations restricted significant followup programs with industry on interesting computer science research ideas developed in the universities and research labs that we normally fund.

We left it to industry on their own to pick up whatever ideas seemed most interesting to them. We feel this laissez-faire approach is no longer appropriate for Defense.

We plan to make a concerted effort to involve industry directly in our plans for building and using future super intelligent computers. Of course, we plan to continue our basic 6.1 research program as a level of effort activity but we intend to substantially augment it by this new 6.2 exploratory development effort.

At the present time, DARPA devotes approximately 10 percent of the total DARPA budget of almost $800 million to microelectronics and computer research. I expect over the next few years to recommend significant increases in funding for these areas since they are so ubiquitous, touching on virtually every aspect of Defense. Other major funding contributors to this area besides Defense are the National Science Foundation, NASA and DOE.

Our relations with these other organizations are excellent. Although interest in basic computer science, hardware and software in particular, remains strong in each of these organizations, the strongest interest in artificial intelligence, other than at DARPA, is at the Office of Naval Research and NSF.

Artificial intelligence is often characterized by its requirement for massive amounts of symbolic processing rather than straight numerical computation as is typically encountered in other scientific disciplines.

Most of the interest in DOE and NASA, however, appears to be of the traditional scientific type associated with the large number-crunchers, such as those you have heard about created by Cray Research, the Control Data Corp. and by Denelcor.

This is an area supported by a handful of computer companies whose domestic business base consists almost exclusively of the

Government and a small number of major oil companies who use computers in geological analyses related to oil exploration.

Although a somewhat broader business base exists in foreign countries, U.S. Government export control restrictions have severely limited access by U.S. companies to this market.

These domestic computer companies struggle to keep afloat in the face of the limited demand. Cooperation within Government to expand the use of these machines and stabilize the Government demand could provide a needed shot in the arm for the U.S. industry.

We plan to work with DOE, NASA and the other organizations involved in funding this area to insure close cooperation. The able leadership of the President's science advisor in this endeavor is most welcome.

Toward that end, a series of meetings were held at the White House under Dr. Keyworth's sponsorship this past spring to discuss the Government's role in supercomputer technology.

It was decided to set up three working groups, two of which were to be chaired by DOE and one by DOD. The DOE topics were how to stimulate supercomputer development, and how to provide wider access to them.

The focus of these two DOE-led groups is on high speed numerical machines. DOD agreed to take the lead in stimulating the exchange of programmatic information on the various Government programs including the new Defense initiatives.

For a number of years, the Navy has been funding development work at Lawrence Livermore National Laboratory on a high performance machine known as the S-1. We understand that funding for this effort will now be shared between DOE and the Navy.

We plan to work closely with the Navy on its S-1 computer development by exploring how it might be used for symbolic processing. Similarly, we hope to explore several related areas jointly with DOE, NASA and NSF. If access to Cray machines becomes easily available, the DOD may find it attractive to become a limited user of these machines as well.

Defense is not a heavy user of numerical supercomputers today. Of the 50-odd supercomputers which have been sold by this small enterprise in the United States over the last few years, I would imagine that there are not more than four that are directly engaged in defense work today.

I wish to now focus my attention on one of the primary issues before the committee; namely, the health of the computer science field. Although no one doubts the immense potential of this area, the field remains badly in need of trained scientists and engineers.

The demands of a few large companies could easily consume all the Ph. D. graduates of this discipline in any given year. The need for training at the masters level and higher is also acute if we are to rapidly buildup the necessary infrastructure within industry.

With industry competing fiercely for faculty members as well as new graduates, the pressures are great for diffusion of academically oriented researchers into industry. This has an added disadvantage of further reducing the residual capability to train new students and to do the most long-range research that is needed in this country.

DARPA will continue to fund university research in computer science with the explicit aim of increasing the numbers of qualified faculty and graduate students over the next few years.

Experimental computer equipment is also badly needed in computer science since the field remains highly experimental. In fact, a range of equipment and instrumentation is needed ranging from individual work stations attached to local communications networks at one end to the largest available machines at the other.

Without adequate facilities, progress will certainly be slow. Both DOD and NSF initiated limited efforts in the early 1980's to provide added computational facilities and access. However, the state-of-the-art moves very rapidly in this field and further efforts will be required to both update and maintain these facilities.

Major non-Federal support for computer science has been provided by some U.S. corporations such as IBM, Xerox, and AT&T. However, these funds are generally not available for university support and have tended to be oriented toward business product development rather than long-range research. Still, the quality of research carried out by these companies remains impressive.

I am particularly gratified by the cooperative efforts indicated by the SRC representative here today, Larry Sumney, who is an old colleague from Defense, and the efforts of the MCC, led by another Defense leader of the recent past, Adm. Bobby Inman, in training to redress some of the shortfall of industry contributions to basic research in these areas.

Recent changes in the tax laws have also made it increasingly attractive for equipment manufacturers to provide hardware to nonprofit institutions, such as universities, without significant penalty.

This has been a very positive impact in stimulating university/industry relations and in making badly needed equipment available.

The new research ideas in the community are still vibrant and alive. We expect significant progress in distributed systems development and software over the next decade along with the appearance of very powerful machines perhaps 3–4 orders of magnitude faster than those generally available today.

Through the use of VLSI, specialized architectures and parallelism, we believe significant performance improvements will be possible. We expect to capture expertise in knowledge bases for use with intelligent machines and we expect to make VLSI technology available to researchers to create whole new systems rapidly and at greatly reduced cost.

By maintaining a strong basic research program while simultaneously reaching out to involve industry, we envision a fundamental strengthening of our industrial base in the country.

This, we believe, will benefit both Defense and the rest of the Nation, but most importantly, we feel that this technology when fully developed will provide the competitive edge over our military adversaries in the decade of the 1990's.

[The prepared statement of Dr. Cooper follows:]

PREPARED STATEMENT OF DR. ROBERT S. COOPER, DIRECTOR, DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

I'm pleased to be here today at your invitation to discuss my view of the state of computer science research in the U.S. and to review briefly our plans for developing super intelligent computers for defense.

As you know, DARPA has been a major supporter of computer science research since the early 1960's, shortly after the agency was created by President Eisenhower in the wake of Sputnik. For twenty years we have played a central role in bringing new technological concepts to fruition for the benefit of defense and also, I might add, creating spinoff benefits for the civilian sector. DARPA is known for pioneering the development of computer time-sharing in the early 1960's, computer graphics in the mid 1960's, and packet switching technology (and the Arpanet in particular) in the late 1960's and early 1970's. During the 1970's we led the development of network protocols, demonstrated the application of packet switching to radio and satellite systems and recently have embarked on an ambitious program to encourage more widespread use of VLSI design and fabrication by our leading universities and research labs, relying on commercial vendors for actual fabrication of integrated circuit chips.

Since our earliest days as an Agency, DARPA (then ARPA) has provided a major portion of the nation's support for Artificial Intelligence research. This work is aimed at developing machines that will have human-like capabilities for intelligent behavior. As you know, this technology is expected to form the basis of the so called "Fifth Generation Computer." We believe this field holds special promise to produce a class of very intelligent computers such as Defense will require within the next few years.

In the FY84 President's Budget submission to the Congress, we included a new program entitled "Strategic Computing & Survivability" which directly addresses this need. Over the next few years, we plan to develop a class of super intelligent machines and to apply them to a wide range of critical defense problems. Three initial areas have been identified as a specific focus for this work; however, we anticipate that others may also prove to be of high priority as the technology base evolves. In particular, a Defense Science Board task force chaired by Professor Joshua Lederberg of Rockefeller

University has been established to study the area of military priorities for applications of super intelligent computers. We expect to have their results early next year.

This program represents somewhat of a departure for DARPA in that up to now, funding limitations restricted significant follow-up programs with industry on interesting computer science research ideas developed in the universities and research labs. We left it to industry, on their own, to pick up whatever ideas seemed most interesting to them. We feel this laissez-faire approach is no longer appropriate for Defense. We plan to make a concerted effort to involve industry directly in our plans for building and using future super intelligent computers. Of course, we plan to continue our basic 6.1 research program as a level of effort activity but we intend to substantially augment it by this new 6.2 exploratory development effort.

At the present time, DARPA devotes approximately 10% of the total DARPA budget of almost $800M to micro-electronics and computer research. I expect over the next few years to recommend significant increases in funding for these areas since they are so ubiquitous, touching on virtually every aspect of Defense. Other major funding contributors to this area besides Defense are the National Science Foundation, NASA and DOE.

Our relations with these other organizations is excellent. Although interest in basic computer science, hardware and software in particular, remains strong in each of these organizations, the strongest interest in Artificial Intelligence (other than DARPA) is at the Office of Naval Research and NSF. AI is often characterized by its requirement for massive amounts of symbolic processing rather than straight numerical computation as is typically encountered in other scientific disciplines.

Most of the interest in DOE and NASA, however, appears to be of the traditional scientific type associated with the large "number-crunchers" such as the CRAY-1, the Control Data Corporation's Cyber series of super computers and by Denelcor. This is an area supported by a handful of computer companies whose domestic business base consists almost exclusively of the Government and a small number of major oil companies who use computers in geological analyses related to oil exploration. Although a somewhat broader business base exists in foreign countries, U.S.

Government export control restrictions have severely limited access by U.S. companies to this market. These domestic computer companies struggle to keep afloat in the face of the limited demand. Cooperation within the Government to expand the use of these machines and stabilize the government demand could provide a needed shot-in-the-arm for the U.S. industry. We plan to work with DOE, NASA and the other organizations involved in funding this area to insure close cooperation. The able leadership of the President's science advisor in this endeavor is most welcome.

Toward that end, a series of meetings were held at the White House under Dr. Keyworth's sponsorship this past spring to discuss the Government's role in "supercomputer" technology. It was decided to set up three working groups, two of which were to be chaired by DOE and one by DoD. The DOE topics were "how to stimulate supercomputer development" and how "to provide wider access to them." The focus of these two DOE led groups is on high speed numerical machines. DoD agreed to take the lead in stimulating the exchange of programmatic information on the various government programs including the new Defense initiatives.

For a number of years, the Navy has been funding development work at Lawrence Livermore National Laboratory on a high performance machine known as the S-1. We understand that funding for this effort will now be shared between DOE and the Navy. We plan to work closely with the Navy on its S-1 computer development by exploring how it might be used for symbolic processing. Similarly, we hope to explore several related areas jointly with DOE, NASA and NSF. If access to CRAY machines becomes easily available, the DoD may find it attractive to become a limited user of these machines as well.

I wish to now focus my attention on one of the primary issues before the committee: namely, the health of the computer science field. Although no one doubts the immense potential of this area, the field remains badly in need of trained scientists and engineers. The demands of a few large companies could easily consume all the Ph.D. graduates of this discipline in any given year. The need for training at the Masters level and higher is also acute if we are to rapidly build up the necessary infrastructure within industry. With industry competing fiercely for faculty members as well as new graduates the pressures are great for diffusion of academically oriented researchers into industry. This has the added disadvantage of further reducing the residual capability to train new students and to do the most long range research. DARPA will continue to fund university research in computer science with the explicit aim of increasing the numbers of qualified faculty and graduate students over the next few years.

Experimental computer equipment is also badly needed in computer science, since the field remains highly experimental. In fact, a range of equipment and instrumentation is needed ranging from individual workstations attached to local communications networks at one end to the largest available machines at the other. Without adequate facilities, progress will certainly be slow. Both DoD and NSF initiated limited efforts in the early 1980's to provide added computational facilities and access. However, the state of the art moves very rapidly in this field and further efforts will be required to both update and maintain these facilities.

Major nonfederal support for computer science has been provided by some U.S. corporations such as IBM, Xerox and ATT. However, these funds are generally not available for university support and have tended to be oriented toward business product development rather than long range research. Still, the quality of research carried out by these companies remains impressive.

Recent changes in the tax laws have also made it increasingly attractive for equipment manufacturers to provide hardware to non-profit institutions without significant penalty. This has had a very positive impact in stimulating university/industry relations and in making badly needed equipment available.

The new research ideas in the community are still vibrant and alive. We expect significant progress in distributed systems development and software over the next decade along with the appearance of very powerful machines perhaps 3-4 orders of magnitude faster than those generally available today. Through the use of VLSI, specialized architectures and parallelism, we believe significant performance improvements will be possible. We expect to capture expertise in knowledge bases for use with intelligent machines and we expect to make VLSI technology available to researchers to create whole new systems rapidly and at greatly reduced cost

By maintaining a strong basic research program while simultaneously reaching out to involve industry, we envision a fundamental strengthening of our industrial base in the country. This we believe will benefit both Defense and the rest of the nation. But most importantly we feel that this technology, when fully developed, will provide the competitive edge over our military adversaries in the decade of the 90's.

Mr. **FUQUA**. Thank you very much.

Dr. Trivelpiece, we are very happy to welcome you back here again.

Mr. **TRIVELPIECE**. Thank you, Mr. Chairman, members of the subcommittee.

To me it is axiomatic that the United States which, in effect, created the supercomputer business should not relinquish its lead in this particular area to anyone. As I listened to the previous panel members speak and heard some of their concerns, I am reminded of the rather tired joke that there are three great lies in the world. One is "I will still love you in the morning." The other one is "The check is in the mail." I believe the third one is "We are from the Government. We are here to help you."

I think you have addressed the proper rhetoric question at the outset; namely, is this an area where the Government should become involved and, if so, how and in what way?

I believe the answer is, yes, and I would like to address that answer now. You have my prepared statement:

I ask that it be submitted for the record.

Mr. **FUQUA**. Yes. Without objection, it will be made part of the record in its entirety.

Mr. **TRIVELPIECE**. Thank you.

I would like to point out that the Atomic Energy Commission and the successive agencies, ERDA and DOE, had a rather pivotal role to play in the creation of supercomputers in the United States. They, in a sense, were friendly purchasers of the latest state-of-the-art technology in the way of hardware as it existed at any particular time, and usually purchased either the first or nearly the first of each generation and contributed to its value greatly by helping develop the software that made that particular collection of apparatus more valuable.

To that extent, I would like to acknowledge here today two somewhat unsung heros of that particular activity. One of them happens to be present in the audience. Sidney Fernbach was at Lawrence Livermore Laboratory and was the individual who made some of these purchases in a way that accelerated the supercomputer industry. The other was John Abbadessa, Controller of the Atomic Energy Commission. During the period he was controller, he fought for, and proposed to incorporate into AEC budgets, the necessary funds to begin some of these generations of supercomputers.

The value of these devices has been that they have permitted exploration of various areas of science. In the case of the Atomic Energy Commission, that contribution was in trying to develop an understanding of the physics and the operational characteristics of nuclear weapons. The benefits that accrued from that have been very substantial.

Other technologies that are of interest to the Government have benefited. I think the fusion program is one that not only benefited, but, in fact, would be seriously handicapped in its ability to make progress in the future if it were not for the kinds of supercomputers that have come into existence.

We are now beginning to discuss the possibility of building multibillion-dollar facilities, and the cost of such facilities is such that they should be designed using the latest available knowledge. The kind of knowledge that is needed is that which will enable us to correctly understand, analyze and predict the characteristics of hot plasmas in order to confine them.

So, given that circumstance, it doesn't seem to me that the computer revolution has even arrived. One of the earlier speakers pointed out that there is a computer revolution. I think that the computer revolution is yet to come, and that when it does, we will find rather startling things. An ability to design various facilities such as advanced fusion systems is going to be incredibly important to us.

In the sense that the supercomputers that we are talking about have cost money, you can say, well, the Government played a role. It paid that price. The Government it has gotten goods and services for value given.

I think that in fact it goes beyond that. The cost of these computers has been negative by an incredible amount. If you look at it in one sense, the Department of Energy has created a technology transfer activity that is, to some extent, rather incredible. It has caused to come into existence a particular computer. It has contributed to the development of software, thereby adding value to it. The vendors have then sold these to other organizations and institutions; the oil companies have benefited from them. The question is what to do next, what should we do in terms of the future.

There has been a tendency in purchasing computers in the past to focus on what is referred to as benchmarking; that is, a particular code is used as a basis for a procurement action in which the need is to meet or beat the competition in terms of having that particular code run on a particular computer faster than the competition.

That, I think, does not address what I perceive to be the pattern of the future. The pattern of the past was to buy all the hardware that you could put in the smallest container and learn how to use it. The pattern of the future is going to involve the simultaneous development of software along with the hardware in an integrated way. It is in developing a plan for that that I think some time is needed because I don't believe everyone has a clear picture of what the future is going to be.

The question is how to create the right collection of conditions in which this future pattern is going to have to be solved, I think a way to consider doing this is to try to define four or five technical problems, scientific problems of major national importance and then make a requirement that the vendors, together with whatever support they choose to define, develop machines that solve those problems, not just simply passing a simple benchmark test and competing in that mode.

This would require them to go outside of the simple hardware aspects of their activities and involve physicists, chemists, engineers, mathematicians and computer scientists capable of putting together the right collection of mathematics in order to solve the problem with the industry, so that the products developed are machines that solve problems.

Later this summer I hope to conduct some workshops to try to engage some of the members of the scientific community in a debate as to what constitutes those problems.

Two of these problems that I have mentioned include some of the weapons applications in which improved computational ability is required, and the fusion program in which an ability to accurately analyze the characteristics of hot plasma, predict their characteristics in a way that will enable you to design future generations of devices is essential and going to be incredibly cost effective as compared with the sums that we are talking about for building these kinds of facilities.

Other problems that are quite familiar to many people are the geophysical problems associated with oil exploration and reservoir management and other mentioned by the previous speakers.

Mr. Chairman, I would like to repeat what I said at the beginning. It is axiomatic to me that the United States should not relinquish its lead in the supercomputer field. The Government does have a proper role to play.

It is going to take some time in planning to determine precisely what that role is. The DOE will certainly cooperate with the Congress in trying to help develop an appropriate plan.

Thank you.

[The prepared statement of Mr. Trivelpiece follows:]

PREPARED STATEMENT OF ALVIN W. TRIVELPIECE, DIRECTOR, OFFICE OF ENERGY RESEARCH, DEPARTMENT OF ENERGY

Mr. Chairman and Members of the Subcommittee:

I am pleased to have the opportunity to testify before you on computers and their role in energy research. My testimony today will focus primarily on the so-called supercomputers. I will discuss some of the past and present usage of supercomputers by the scientists and engineers working for the Department of Energy's contractors. Then, I will turn to questions of the future of large-scale scientific computing that you raised in your letter of May 24, 1983.

Supercomputers are the fastest general purpose computers now available to solve large-scale scientific problems. Today, they are capable of performing 10's to 100's of millions of arithmetic operations per second and have several million words of internal memory. Scientific computing and supercomputers are quite different from artificial intelligence applications of computers and the fast computing engines that may be required for those applications. My testimony will be devoted to the scientific applications of supercomputers.

Introduction

The Supercomputer industry has been the exclusive domain of U.S. vendors up to the present. Computer work done at U.S. laboratories and universities to develop the software required to effectively use the new machines developed by vendors has resulted in a variety of scientific computer applications. These applications, or programs, are made available to industry at a nominal cost through several distribution organizations. Some of those applications are, in fact, adopted by industry, particularly when the techniques are demonstrated to be valuable in modeling complex processes.

Examples of current applications in use by industry are:

Petroleum Industry: Analysis of vast amounts of seismic data to locate promising sites for drilling; reservoir modeling for tertiary recovery.

Automotive Industry: combustion modeling for internal combustion engines using flame front tracking and chemical kinetics programs; computer aided design and testing of automobile parts.

Electronics Industry: computer aided design of very large systems integrated circuits; numerical solution of large sets of differential equations describing circuits using numerical linear algebra techniques.

Manufacturing: computer aided design of machine parts; testing structural stress-strain properties of structures using finite element methods; non-destructive evaluation using image reconstruction techniques.

Power Industry: modeling of nuclear reactor safety systems; power grid optimization using linear programming techniques.

Chemical Industry: process control and optimization; simulation of polymers; reaction diffusion modeling and turbulence.

Movie Industry: computer graphics techniques for image generation and animation.

Clearly, supercomputers are having a great impact on many high technology activities in our country and maintaining U.S. leadership in the future is vitally important.

To maintain that leadership will require not only the development of advanced computer technology but a re-emphasis on the close collaboration of the Federal government, industry, and universities that existed during the early years of computer development.

The Federal role in the development of large-scale scientific computers has been a beneficial one. The first electronic computer, the ENIAC, was built by the University of Pennsylvania under contract to the U.S. Army Ballistics Research Laboratory, and the first program run of the ENIAC was a nuclear weapons design calculation by the Los Alamos Scientific Laboratory. In fact, the Manhattan Project and the ENIAC project were initiated the same year (1943). In the succeeding four decades, the defense and computer communities have had a mutually beneficial relationship: nuclear weapons programs have provided much of the motivation and funding for the development of supercomputers, and the computer industry has provided the tools for solving the highly complex problems that otherwise would have remained intractable.

The speed of computation achieved by the fastest computers has increased by a factor of 10 million from the early 1940's to the present. An increase of one thousand resulted from the transition from electromechanical to electronic circuits. Another factor of one thousand has been achieved in the speed of the electronic components (for example, the Cray-1 and the Cyber-205 are 1000 times faster than the computers of the early 1950's). Another factor of ten in this overall increase has come about through increasing levels of parallel execution of computational tasks.

The supercomputer industry is now facing a demand for speed improvements of at least two hundred times in this decade. Unfortunately, the projected speed increases available from faster components appear to be limited to at most about ten during the next decade. Hence, it is clear that new (and explicit) forms of parallel computer organization will be required to meet the needs of the nation. Achieving success in this radical change in the design of supercomputers will require some of the best minds in the nation.

The early years of the history of supercomputers were characterized by a strong interaction among government, universities, and industry that guided and accelerated the development of very powerful scientific computers. The mathematician, John von Neumann, served as an important catalyst for this interaction: he was a consultant to both the government and the computer industry, and he led the development of the computer at the Institute for Advanced Studies at Princeton that served as a model for the development of a whole generation of computers, including the IBM 700 series computers and their successors, the ILLIAC, the Johnniac, the MANIAC, the ORDVAC, and the ORACLE.

This highly beneficial interaction among government, universities, and industry continued through the decades of the 1950's and 1960's but suffered a serious decline in the 1970's. Universities, in particular, have become somewhat isolated from the main stream of supercomputer research and development. The leading supercomputer of the early 1970's was the Control Data Corporation 7600, but none of these computers was installed at an American university (although several were installed at foreign universities). Even today, only three American universities have on-site supercomputers. This unfortunate situation has resulted in fewer graduates with backgrounds in large-scale scientific computing and fewer research projects being conducted in universities in this field than in previous decades. There is general agreement among these three communities that it is essential to reestablish their close collaboration of earlier decades for the United States to maintain its leadership in supercomputer development, marketing, and applications.

DOE Supercomputer Applications

The domination of supercomputer industry by U.S. vendors has contributed to the ability of the U.S. government laboratories involved in defense, space, aeronautics, weather forcasting etc., to maintain their leadership position in these fields. In turn, the DOE laboratories have played an important role in causing the development of each new generation of supercomputer systems. DOE's weapons laboratories have historically required the use of the fastest computers available in order to carry out their missions. Since the 1950's, a DOE laboratory has acquired the first, or one of the first, manufactured units of nearly every new large-scale computer. These early models would arrive almost devoid of usable software and, consequently, the laboratories were required to provide the necessary systems and application programs in order to effectively utilize the new machines.

Every generation of large scientific computers has had a significant impact on the ability of the DOE laboratories to meet their mission goals. Each new machine has improved the laboratories' abilities to more accurately model complex physical phenomena, thus improving their productive capability and reducing dependence on empirical scaling. This is particularly true in the design and testing of nuclear weapons and more recently in controlled nuclear fusion research. Approximately 35% of the total number of supercomputers currently in use in the U.S. are located in DOE laboratories.

The highly successful nuclear weapons program of the U.S. has depended critically on supercomputers to maintain its leadership. Early weapons designs were heavily dependent on field testing, but this approach is time consuming, costly, and sometimes constrained by other factors. The mixture of computation and field testing has shifted significantly from primary dependence on testing to a much higher dependence on computation in the design process, resulting in very sophisticated designs that have been produced on shorter time scales and at lower costs than could have been achieved without the aid of supercomputers. In this sense, the cost of developing the computers has been negative, i.e., they saved in tests more than they cost. As nuclear weapons have become increasingly complex, the design calculations have also become increasingly complex, and further progress in nuclear weapons design will be paced and perhaps limited by the speed of the computers available.

Two other DOE programs that have utilized supercomputers extensively are Magnetic Fusion Energy and Inertial Confinement Fusion. In both programs, supercomputers are required to model the complex behavior of hot plasmas including the effects of electric and magnetic fields, atomic physics, the interaction with intense radiation, and various boundary conditions. Simulations of thermonuclear plasmas have been used extensively to interpret experimental data and to design future experiments. For example, new experimental devices in the magnetic fusion program now cost several 10's to several 100's of millions of dollars and require 2-7 years to build. The fusion program cannot afford to use an empirical approach to design its devices, since this would be too costly and time consuming.

Large computer codes are used extensively in the complicated process of optimizing magnetic field design to maximize plasma pressure and to minimize energy transport, to determine heating system requirements, and for electrical and mechanical design. While substantial success has been achieved in the modeling used to design fusion experiments, even greater advances will be required to achieve the accurate design information which will be required for future fusion test reactors and eventually for commercial power plants which will likely cost several billion dollars each. This increased accuracy will come from improved plasma models and the substantially increased computing power that will be obtained from future generations of supercomputers.

The magnetic fusion program has been unique in making its supercomputers accessible to all of its contractors in the national laboratories, universities, and industry. This has been accomplished through the Magnetic Fusion Energy Computer Network that was initiated in 1974. This network has its center at the Lawrence Livermore National Laboratory where a CDC 7600 and two Cray-1 computers provide large-scale computing capability to the whole program. Satellite communications are used to link the four largest fusion program contractors (Princeton Plasma Physics Laboratory, Oak Ridge National Laboratory, Los Alamos National Laboratory, and General Atomic) outside of Livermore to these computers. Other contractors have telephone line access. This network has proven to be the best way to provide cost-effective, large-scale computing to a large group of geographically dispersed users. The network approach has allowed the fusion program to match computing resources to the needs of all of its contractors at once rather than on a laboratory by laboratory basis.

Code sharing is greatly facilitated since all contractors use the same computer system. This has helped eliminate duplication of the development of large, expensive applications codes. In addition, the network has substantially increased communication and cooperation among fusion contractors.

Although the Department's laboratories have a relatively large number of supercomputers, most researchers supported by the Office of Energy Research, in High Energy and Nuclear Physics, Basic Energy Sciences, and Health and Environmental Research do not have access to these computers which are fully utilized by Defense Programs and Magnetic Fusion. A recent study conducted for the Office of Energy Research indicates that in some areas its researchers are giving up their traditional position at the leading edge because of their lack of opportunity to do large-scale computations. This study further indicates that there is currently a need for one machine of the present generation of supercomputers to satisfy the requirements of the programs mentioned above. A promising solution for meeting this need is to serve this geographically dispersed group of researchers by a computer network similar to the one in the magnetic fusion program. This would allow university contractors as well as national laboratories to have access. As an interim solution to this problem, beginning in early May of this year, 5% of the Cray-1 time on the Magnetic Fusion Network was made available to the contractors of other programs in Energy Research. This represents some hardship to the fusion program, which itself is short of computing capacity. I am currently considering options for a permanent solution for an Energy Research computer network.

In making supercomputer time available to its university contractors, the Department is also helping to solve the nationwide problem mentioned earlier. At the present time, very few universities have adequate access to such large-scale scientific computers, and consequently, too few university students are trained in their use. In addition, if more university personnel had access to supercomputers, the enormous power of these machines would bring substantial advances to many new areas of research and engineering.

Future Generations of Supercomputers

I would like to turn now to the important question of development of future generations of supercomputers and the role that the Department of Energy should play. Several months ago, I convened an internal DOE committee to examine this issue. This committee has consulted with both users and technical experts in our laboratories and universities. The committee has also had discussions with the supercomputer vendors on possible government roles in the development of future systems. First, I will address the near-term problem of this decade and then I will discuss longer-term research.

Until the late 1950's, a new computer meant an order of magnitude (10x) increase in capability. Since that time, however, the increase has been more like four-fold (4x) with each new major mainframe. At the same time, the computational needs of science and technology programs have been increasing even more rapidly. Consequently, supercomputer technology has fallen short of meeting modern computing requirements.

Today, the national laboratories have major programs that require a 200 times increase in capability by the end of this decade. If the development pattern of the past were to continue, supercomputer systems would deliver far less capability than our scientists and engineers could effectively use.

The Japanese, recognizing a similar need, have formally created a national program of their own to achieve 100 times present capability by 1990, and thereby establish their leadership in supercomputers. The French have also initiated a supercomputer development effort. They recognize the importance of this powerful scientific tool not only in research but in designing and engineering new products for the marketplace.

Nippon Electric Corporation has announced a new machine scheduled for early 1985 with a performance of 1,300 million floating point operations per second. This will be the first time a supercomputer's computation speed has exceeded 1,000 million floating point operations per second.

The U.S supercomputer industry is relatively small. The reason is that the development costs are large and the risks are high. The market is also small, smaller in fact than the market seen by the Japanese competitors, because of our export controls on this technology. For all these reasons, our largest computer manufacturers have chosen not to develop these large machines. Only two companies, Cray Research and Control Data, have produced the latest generation of supercomputers. A third company, Denelcor, has entered the field. Nonetheless, the combination of U.S. companies producing hardware and government laboratories producing much of the software has kept the U.S. in its leadership position. I believe the U.S. should retain its leadership in supercomputers and not relinquish it to non-U.S. firms. The supercomputer vendors, the government, the government laboratories and the universities should play their appropriate roles as they have in the past. Strong efforts by all of these will be necessary for success.

There is continuous interaction between government laboratories and supercomputer vendors. One purpose served by such interactions is to keep the vendors informed of the users' problems. Taking the users needs into consideration, the designs to be used in supercomputers that will be produced during this decade have already been chosen by the vendors. The required performance increases can no longer be obtained by faster components alone and all vendors will be utilizing parallel processors along with faster logic and memory circuits.

Implementation of the new parallel architectures in faster circuit technology to achieve the 200x improvement in performance goal stated earlier will be a tough challenge. However, U.S. vendors have done an excellent job of successfully assembling hardware and making it work in the past and can be expected to do so in the future.

One of the most significant problems in the hardware area is integrated circuits where high quality fast logic and memory circuits are required. Again, there is a market problem because the supercomputer industry buys relatively few circuits compared with the large number used for consumer products, for example. The U.S. supercomputer industry already buys substantial numbers of circuits from Japan. However, this raises a concern. Will U.S. vendors be able to obtain the most advanced circuits on a timely basis from foreign vendors? I believe that this is a high risk approach and the government should fund research in the basic underlying technologies to improve U.S. competitiveness.

In the past, the realization of the full potential of supercomputer hardware as delivered by the vendor has been limited by software.

As stated earlier, supercomputer vendors have provided little software to the user and the Department's laboratories at Livermore and Los Alamos have had to develop substantial amounts of software for the previous generations of machines. Since the development of software to utilize the new parallel processor machines will be a greater challenge than has been faced with previous generations of machines, it is important that we get an early start and that we involve the most capable institutions in the problems. Specifically, some of the Department's laboratories have extensive experience in developing operating systems and algorithms. Universities also have excellent capabilities to develop advanced algorithms. I believe that a sensible approach to the software problem is to place several computers with interesting parallel architectures for research purposes in some of the Department's laboratories as soon as possible, and make them available to our contractors through a communications network. In fact, the Department has already taken a step in that direction by obtaining a Denelcor Heterogeneous Element Processor machine for Los Alamos. It should be noted that there is a clear distinction between this type of "experimental" machine and computers which are purchased to do production jobs. I believe that the Department should also buy early prototype machines of each new generation of supercomputers and make them available to a wide community of users in order to accelerate the software development process. In addition to aiding software development, these purchases will stimulate the industry financially and encourage much needed internal company research. At all phases of the software development there will have to be a continuous interaction with the supercomputer vendors. This tight coupling will provide important feedback to the vendor on his product.

The actions recommended above are some steps I believe will accelerate supercomputer development in this decade. For computers with capabilities beyond the 100-200x present machines that we expect this decade, new approaches other than those currently planned by vendors may be necessary. Long-range research in new architectures, mostly in universities, should be enhanced by the Department of Energy and other parts of the government. One possible means for choosing the most promising approaches is to have a national competition. The competition would be based on specifying the problems that we need to have solved and then encouraging companies to offer to sell the Government machines that will do the job. This would allow for maximum flexibility on the part of the vendors. The machines would, no doubt, be expensive; but the money would be better focused than it would be if we were to start up an industry-laboratory-university program to develop and design supercomputers to some artificial set of specifications.

In summary, supercomputers are essential to maintaining the country's leadership in defense technologies, many areas of research and development, and design and engineering of new commercial products. Supercomputers have become such an indispensible scientific and engineering tool that the U.S. cannot afford to relinquish its lead in this key industry. The U.S. can maintain its leadership, but it will take an extensive effort.

Mr. Chairman, this concludes my prepared statement. I would be pleased to answer any questions at this time.

Mr. Fuqua. Thank you very much, Dr. Trivelpiece.

Dr. Kerrebrock, we would be pleased to hear from you at this time.

Mr. Kerrebrock. Chairman Fuqua, Chairman Lloyd, and members of the subcommittee, I am happy to present to you today the current NASA capability and our plans for supercomputers and their possible application to energy research. NASA currently has class VI computational capability at the NASA centers. Ames Research Center has had access since 1982 to a Cray IS which will be upgraded to a Cray XMP in the fall of 1983. This system will support primarily computational aerodynamic research. The Lewis Research Center recently installed a Cray IS computer system with 2 million words of main memory.

The system will be used primarily for internal computational fluid mechanics, and for thermal and structural performance analysis of propulsion system components.

A CYBER 205 now operational at Goddard Space Flight Center is being used for atmospheric science investigation.

The centerpiece of our proposed aeronautics program for fiscal year 1984 is the new numerical aerodynamic simulation (NAS) capability, which will be developed over the course of the next few years.

The NAS capability will include a large computer system which will make it possible to perform most of the calculations required to design a new aircraft with sufficient accuracy and reliability to eventually reduce the long and expensive process of wind tunnel and flight testing now necessary.

There is every reason to believe that the computer technology to achieve this objective can be developed in the next few years. The NASA commitment to NAS will provide incentives for various elements of the computer industry to produce machines which will be able to take advantage of major commercial applications. Developments of this kind will enhance the competitive posture of the United States in this important international market.

The NAS program will also serve as a powerful pathfinder in basic research. Solutions of the equations of fluid mechanics will reveal underlying mechanisms of turbulence, flow separation and reattachment, and aerodynamic noise.

This understanding will contribute to the design of quieter and more efficient engines and vehicles required for continued U.S. aerospace pre-emence. The NAS capability will allow similar advances in other computationally dependent disciplines such as structures and structural dynamics, heat transfer, chemistry, and weather and climate modeling.

Our basic commitment is to maintain NASA in a position of international leadership in computational fluid dynamics, and the schedule is paced to take advantage of advances in computer science within U.S. industry.

This intent is most directly expressed in terms of the expected speed and memory of the high speed processor. Initially, in 1984, the available computational speed is planned to be 250 million floating point operations per second with 64 million words of high speed memory.

By 1987, this capability is planned to increase to 1 billion floating point operations per second, with 256 million words of high speed memory. Beyond 1987, as more advanced high speed processors become available they will be incorporated into the NAS system.

During the 1980's, NASA will orbit sensors that can generate data at rates up to 10 to the 13th power (10^{13}) bits per day. A variety of complex image processing tasks such as geometric correction, correlation, image registration, feature extraction, and multispectral classification are required to extract useful information from space-acquired data. NASA's space research computational workload is expected to increase from 1 to 10 billion operations per second.

To meet this need, NASA initiated a research and technology program to evaluate the use of a massively parallel computer architecture. This concept is based on the use of 16,000 processors operating simultaneously to achieve the needed speed. Under a NASA contract awarded in 1979, the massively parallel processor (MPP) has been developed. The heart of this computer is a custom integrated circuit containing eight complete processors. Integration, test, and delivery of the MPP was completed in May of 1983 and processing research on space-acquired information using the MPP will be initiated this summer.

The MPP, designed to meet a specific NASA need, is a current example of the successful development of a special purpose supercomputer. Research in alternative computer architectures will continue as part of the OAST computer science program with future development of special purpose computers emerging in response to NASA needs and research opportunities.

The committee asked about the possible applications of supercomputers to NASA's work in support of the Department of Energy. The Lewis Research Center is looking into the use of its Cray 1 computer to study the structural dynamic response of large wind turbine blades in turbulent wind conditions.

Should this capability appear useful in support of the MOD-5 wind turbine project, LERC will propose to DOE that we proceed along this line of investigation. There are many other possibilities.

Certainly, computational fluid dynamics could be applied in aerodynamic research for large wind turbines employing advanced blade geometries and concepts for wind turbine controls, such as ailerons, to analytically study multidimensional airflow around aileron controlled, rotating blades operating at or near stall conditions.

There will be applications to analysis of three dimensional flow in small compressors and turbines such as are needed for automative gas turbines. Should the use of these or other advanced computational techniques appear to provide benefits to current or future DOE projects we support, we will bring the matter to the attention of DOE for its consideration.

This completes my formal testimony at this point. I would like to thank you for the opportunity to present it.

I would like to add a personal note, which is my opinion on the most attractive role of the Government in support of supercomputers. I believe NASA should have two distinguishable roles in furthering the development and utilization of supercomputers in the United States.

First, I think we should encourage private development of advanced supercomputers by acting as an aggressive and supportive first user and, second, by conducting supporting research in computer science and technology particular to NASA needs.

These roles are currently recognized by Government agencies. However, increased emphasis is warranted in light of the current technical opportunities and international competitive situations.

I think that this view coincides fairly closely with those that were expressed by the earlier panel. By making full use of industrial prototype supercomputers, Federal R&D agencies can assure early evaluation of technical opportunities and assist in the early recovery of private R&D investments.

The prototype hardware and software will, of course, enhance the user agency's capabilities in R&D and its earliest possible use will stimulate beneficial interaction between the user community and the developer.

Enhanced capability to address many technical problems of critical importance will result as in the example of [CFD] computational fluid dynamics which I have described. Since the development costs for prototype hardware [illegible] systems software are low, compared to those in m[illegible] high technology industries, such as the aircraft indu[illegible] rapid recovery of a portion of the development c[illegible]ough early sales to Government R&D agencies can stimulate the rate at which new computers can be introduced.

Finally, because of the high initial software conversion costs to commercial users of new supercomputer products, it is essential that the Government work with the manufacturers to mature prototype designs into commercial systems.

With that, I will stop and answer any questions.

Mr. FUQUA. Thank you very much, Dr. Kerrebrock.

The DOE and DOD and NASA interest in supercomputers, are they compatible or do they overlap?

Dr. COOPER. Let me speak for Defense. It is my view that Defense has little interest in the current fastest computers available, that is, the supercomputers for number-crunching. This is because of the fact that Defense has only a handful of potential applications for computers in that class, that is computers that cost $10, $20 or more million apiece. There are only a few military imperatives that are strong enough to induce individuals within the Department of Defense, which is not monolithic, to go out and buy them and use them for that purpose.

Typical of the uses that we put those computers to are numerical simulation of the atmosphere for weather forecasting at Fleet Weather Center, the NSA needs for decryption of encoded signals and typically the use of computers to understand underground nuclear testing and for understanding weapon effects and that sort of thing.

Aside from these few rather specialized applications, there is no general application in Defense of these machines. Now, that doesn't mean that there couldn't be a broader application.

As a matter of fact, my view of the industry—the industry being the few companies, two or three, who actually build machines of this size—is that they have had a relationship with the Department of Energy which has made them very naive with respect to marketing their product, particularly in Defense. We have had discussions with these people, with CDC and with Cray, never with Denelcor, about how you go about marketing a computer of that size within the Department of Defense. There are alternative applications there, I am sure, but nothing that would stimulate an order of magnitude larger production quantity than is available in that industry today.

So as far as general application of these large number-crunchers in Defense, there isn't much of an application today. Now, that is different from both NASA and DOE, although, again, the size of the market that is there in the Government, in these two organizations, is still small by comparison to the markets that are available to the big computer companies, to IBM, to Digital Equipment Corp., and the others.

These people deal in tens of thousands, hundreds of thousands of units, and we are talking about a few tens of units of these larger computers.

If you talk about the Government's attitude toward these three small companies, you know, the Government is a giant monolith compared to these three little companies, and I am sure that they are over-awed by coming before an august group like this or even trying to deal with something as unspecific as the Department of Defense in a marketing operation.

IBM knows how to market computers for Defense and Defense buys thousands of IBM computers to do its jobs. They know how to market computers.

These small computer companies do not know and understand how to get into Defense and market their computers, so they are relatively ill-equipped to try and develop new markets in that area generally.

Mr. **Fuqua**. Dr. Trivelpiece?

Mr. **Trivelpiece**. There are two classes of problems. One is the data glut problem that exists in a large number of areas. Some of these are ones which Dr. Cooper talked about, where you have to handle massive quantities of data and then be able to sort through it and find intelligent answers quickly. The artificial intelligence application there is very clear.

The other side of that is the collection of scientific problems in which it isn't simply handling the data but, in fact, making progress in areas of mathematics and physics to obtain a better understanding of nature or some particular technical application. In this case, you do require the kind of supercomputer and subsequent generations that we are looking for in order to solve the problems.

You asked the question of whether or not our needs are to some extent compatible. I think they are. The kinds of scientific problems that are coming up do generate data in a glut mode includes biology and medicine, I think even law is generating somewhat of a glut these days in data of the kind that will require computers to handle.

We will benefit from that. I am not sure that I agree entirely with Dr. Cooper, in that there are problems associated with Defense-related activities that are of a scientific or quasi-scientific character for which the kinds of computers we are talking about would be of substantial value.

Perhaps they should be getting involved more than they have been in the past, and that would be to the cumulative benefit of the Nation.

Mr. **Fuqua**. Dr. Kerrebrock.

Dr. **Kerrebrock**. In contrast to what Dr. Cooper said, NASA has an absolute need for staying at the state of the art in high speed computation. This derives from our commitment to excellence in aeronautics, indeed, to world leadership in aeronautics and from the fact that high speed computation a fluid dynamics is the pacing discipline in aeronautics.

It is clear that for the Agency and the United States to maintain leadership in aeronautics, we have to be first in high speed computation for CFD.

In that particular application, brute speed of computation and size of high speed memory are the controlling factors, so for the foreseeable future we can predict a requirement for the best supercomputers that we can lay our hands on.

Our NAS program reflects that commitment. Now, in addition, NASA, in its space activities is information handling limited already. This is true of, for example, the Landsat satellite. It will probably be even more true for the space station when it becomes a reality. We see a parallel need, therefore, for parallel processing of very, very large amounts of data.

One thing that I neglected to say earlier and which I would like to comment on is the university access to supercomputers and our approach to it, which is to establish within NASA a high speed or wide band width satellite link which will tie all of our research centers together and give them access to all of the supercomputers in the NASA research centers, Our hope is that we can give the university community access to that wide band link as well so that to some extent they will have the same access to the NAS capability as the research centers do.

Mr. **Fuqua**. How can the Federal Government, in the development of supercomputers, coordinate these activities between the three agencies that you three represent?

Should it be OSTP or everybody go on their own or is there some coordination needed?

Mr. **Trivelpiece**. Yes. Some coordination is needed, and, at the moment, OSTP has stepped in to fill that need.

I perceive that what they have done and are planning to do is appropriate. I think the assignments that have been made are to go back to our respective agencies and try to develop some plans and incorporate some of those plans in our budget planning process for the forthcoming year. Of course, when they become part of the President's budget, you would hear about them. I think there is a proper role, and at the moment it is being fulfilled by OSTP. That may not persist indefinitely, but it is so at the moment.

Dr. **Cooper**. I agree that a coordinating function is necessary. You have to realize that expanding the use of supercomputers in these three Government agencies and coordinating, trying to stabilize the yearly market for the companies who supply this relatively limited market is not going to be an easy thing to do, particularly because of the fact that within Defense the procurement of computers is done on an individual basis by a wide variety of organizations, and it is the competitive imperative inside of the Defense budget which determines what kinds of computing machines if, indeed, any computing machines are bought.

That and the other single largest influence, I guess, is Representative Brooks and his constraints on the procurement of such machines.

Mr. **Fuqua**. Dr. Cooper, what percentage of DOD R&D funds or DARPA funds are devoted to generic or basic research activities?

Dr. **Cooper**. We have about $108 million in our basic research program this year, but we do research that is closely related to basic research with a substantial portion of the technology part of our budget.

I would suspect that all told, basic research is probably supported to the extent of about 25 percent of the DARPA budget.

Mr. **Fuqua**. Dr. Trivelpiece, to what extent has limited access to supercomputers hindered basic research?

Are the researchers being cut out of an opportunity because of the inadequate access that they have to these computers?

I am really thinking of universities.

Mr. **Trivelpiece**. Mr. Chairman, it is my view that there has been a serious deficiency in access to supercomputers by both students and faculty at most of the major universities in the United States.

Mr. **Fuqua**. They do have access at some of the national labs, and many of them work there, particularly DOE national labs.

Mr. **Trivelpiece**. There are three problems. One is the Nation's computer problem, the supercomputers we are talking about today. Another is the university computing problem. That deals with the present generation of undergraduate students and graduate students learning to manipulate codes on major supercomputers. The other is what I call the ER computing problem. As you know, ER consists of four outlay programs, one of which is magnetic fusion. Fusion has its own computer network which has worked well. The other three outlay programs have not, and in order to try to redress that problem this year, I have caused 5 percent of the time available on the fusion network to be made available to scientists in the other programs.

We had a competitive proposal submission that we reviewed, and now people are taking advantage of it. To some extent that collection of scientists is in part in universities, so we are making a minor contribution toward solving the university computer problem by so doing. However, this could be expanded and with Congress help in 1984, perhaps we can make some expansion of that to include additional universities.

We do have the constraint, however, that we are dealing with ER contractors, so it is those universities which have Energy Research contracts and students that are employed on those contracts. However, that is not a trivial number of students and it will go some distance toward helping alleviate access to systems for graduate students in the United States.

Yes, it is a serious problem. We should give attention to it. Anything that we can do to help in any direction to alleviate it, I think, is to the Nation's benefit.

Mr. **Fuqua**. Ms. Lloyd.

Dr. **Kerrebrock**. I wonder if I might comment on that question.

Mr. **Fuqua**. Certainly.

Dr. **Kerrebrock**. I think we need to rather carefully distinguish two kinds of computing activities in the universities. First of all, let me say that the universities have an almost insatiable appetite for computer time, as does almost any other scientific organization, but there are really two kinds of activities represented in that demand.

One is a bulk computing load. The other is the computing which can only be done on the fastest and largest supercomputer that is available. At any time in the United States there will be one of those somewhere and there will be someone in a university who can only do his research in full competition with the people in the national laboratories if he has access to that most powerful, largest memory computer.

I think that we need to give serious concern to providing access for that class of users. It is that which I was addressing when I mentioned access of universities to our NAS system. We want to make sure that the person out in the university who is pushing the forefront of numerical fluid mechanics has access to the machines as well as the people in the research centers.

Dr. **Cooper**. I agree with that, sir, and I wanted to say that DARPA has no funding now that generally makes available class VI computers for our university sponsored research programs.

All that computational power is on less powerful machines. A substantial number of them, and we do provide very broad access within the computer science community to computing machinery, but none of it is on the class VI machines.

Mr. **Fuqua**. Ms. Lloyd.

Ms. **Lloyd** [presiding]. Thank you, Mr. Chairman. Thank you, gentlemen, for your excellent testimony today. You certainly added much to our hearings.

The Lax report basically stated that we need to do more work in designing and development and advanced supercomputers, that we need greater access to the computers, and that we need to train more scientific minds for the use of computers as well.

How do you see an interface in these objectives with your three respective agencies?

I want to move to the objectives of the Lax report. How do you see your agencies moving in response to this Lax report?

Dr. **Cooper**. From the standpoint of Defense, Defense is not doing anything specifically in any of its programs that respond to any of the three initiatives, although I do recognize the potential desirability of having a broader base of students coming out of our university with familiarity with classic computers and realize that that could have some benefit for Defense in the longer term.

We have set up no specific programs to try and either create the resources to accomplish that or develop specific programs for our computer science support

In the area of development activities, we have none.

Ms. **Lloyd**. I understand that in the 1984 budget that basically that DARPA will have almost a 50-percent increase in the computer area. Certainly it is to our defense peril if we have second rate computers.

Would you address that?

Dr. **Cooper**. Yes, that is true. The thrust of our program is along the lines that I outlined in my testimony, which is largely to create a different kind of computing machinery.

It is to create computers that are able to solve specific kinds of military problems that are not particularly amenable to solution by these large number crunchers.

In particular, they will be machines that process symbols as opposed to numbers with great facility, and will have enormous memories, that is, address space and access to random memories will be key to the functioning of these machines.

They will be closely related to artificial intelligence technologies, and although it is possible to do experimentation on the class VI and larger machines coming on in this general area, we believe that specialized parallel processing architecture involving hundreds, perhaps thousands of microprocessors working with enormous memories will be required to meet the military need.

So our research programs have largely been aimed at this——

Ms. **Lloyd**. Specific categories of problems?

Dr. **Cooper**. And the kind of architectures and computers that we think will be most applicable to those problems for military systems.

Dr. **Kerrebrock**. In addition to our initiative in NAS, which I have probably explained adequately, we in NASA have two special efforts underway to increase our involvement in computer science and in particular with the universities.

We have for some time had what is called the Institute for Computer Applications in Science and Engineering in conjunction with the Langley Research Center, and we have started a parallel to that, the Research Institute for Advanced Computer Science out on the west coast in conjunction with the Ames Research Center.

Now, the intent of both of these organizations is to bring the academic community into close working relationship with our research centers in order to make sure that the advancements in computer science which are coming out of the universities are fully utilized and appreciated by NASA, and that NASA supports the universities.

The other thing we have done is to establish a Center of Emphasis, sometimes called a Center of Excellence, at each of the two major universities, the express intent of which is to assure that the young people who have a deep interest both in computer science and its applications in aerospace technology have an opportunity to develop those in parallel and to interface with the NASA system.

So we are establishing one of these centers at Stanford and one at MIT. Beyond that, we intend to continue to increase the level of activity in computer science and its aerospace applications within NASA.

Ms. **Lloyd**. Dr. Trivelpiece?

Mr. **Trivelpiece**. Each of us represents mission agencies and to some extent the mission of our agency constrains what we can do.

The programs of, say, high energy physics and nuclear physics and so on do offer opportunities for doing research in a way that requires the use of advanced computers and at the same time trains students. So to the extent that we expand the computational capability to address problems associated with these fields, then we do help in the training of students and the advancement of computer sciences and the like.

We also do have a program that is directly related to computer science. I think it has about $15 million in support in fiscal year 1983.

We may well be asking for increased support for that in the forthcoming year. It is a little early in the year to tip our hand.

Ms. **Lloyd**. We will expect to hear from you.

Mr. **Trivelpiece**. We help to train students directly and indirectly and in connection with the mission at the Department of Energy, but we don't have in effect a license to go out and directly support educational institutions in their educational efforts to develop computer science as a field. I think that is still properly the province of the National Science Foundation. But to the extent that our scientific programs caused this to happen, I think we are playing a very important role in improving the circumstances.

Ms. **Lloyd**. Dr. Cooper?

Dr. **Cooper**. I was just going to state that there seems to be a hole in our support for computer science activities in universities from the Government standpoint here.

Typically, the largest supporter of computer science in the Government is DARPA. The National Science Foundation is probably next, followed by the Department of Energy and NASA along with the three military services.

I think the problem is that there has been no real identified need for class VI computers in Defense research, and consequently we haven't—we, the largest funders of universities in computer activities generally, have found no great imperative to go ahead and provide the resources necessary for students to have access to that class of computers.

Now that has changed throughout the years. As a matter of fact, in the early years of the development of these large-scale computers when the IBM 704 was the supercomputer of the day, or when the 7090 was the supercomputer of the day, or when the 7094 stretch was the supercomputer of the day, IBM provided those computers to schools like MIT, where I went to school, and I had access to those supercomputers of that day and was able to work problems.

When IBM abandoned the field of supercomputers, it no longer could make supercomputers available to the universities, and consequently the small companies that now are in that field do not have the ability to put these loss leaders into the university community to build a market for the future, which was what IBM did.

So there is no Federal support at a substantial level to provide those resources today, and there is no ability by this small supercomputer industry to provide those machines as a loss leader with the hope of getting future business.

So there is something wrong with the process there. It certainly has changed since the sixties and early seventies.

Ms. **Lloyd**. The last panel focused on one of the objectives that the Federal Government be more involved in supplying computers for our universities and for our national laboratories as well. Do we have the qualified people to use the computer to make them a viable investment?

Would any of you care to comment on that?

Dr. **Kerrebrock**. I would, if I may, having recently been at a university and expecting to return to one.

The answer is an unequivocal yes, that the universities have talent which will make use of whatever computational capabilities we can provide to them.

In fact, I think that most people would agree that much of the leadership in the utilization of supercomputers does come from the universities. So there is no doubt that they have a legitimate need for this, and that maximization of progress in the United States in the technology of supercomputers and their use depends on making access to these machines available to these people in universities.

Mr. **Trivelpiece**. I would like to quibble with the people on the previous panel to the extent that they talked about making computers available on the campus.

It isn't entirely clear to me that there shouldn't be a very careful distinction drawn between making access available versus managing of computer centers, and it is not entirely clear to me that managing of computer centers is in the university's best interest.

It may well be that making access to supercomputers available to university campuses, given the nature of communication systems today wouldn't be a more desirable end result, so that perhaps some type of regional center such as Congressman Lewis talked about would have some substantial benefits in that it would aggregate the management costs into one location rather than having them distributed across a large number of universities.

Ms. **Lloyd**. You then would agree that we certainly would—that the Federal Government could contribute to greater promise in the computer field by providing access to computer centers?

Dr. **Kerrebrock**. I agree with Dr. Trivelpiece completely on that. There are very few universities that can muster the effort that is required to actually have a class VI computer in their institution and keep it up to date, which is even more important.

Dr. **Cooper**. It is a major management problem to keep control of one of those computers and provide the ready access to it that is required. So it is a major problem that universities are probably not well suited to. But, if you say yes, I am going to make access at a particular location, say it is at the National Center for Atmospheric Research or at one of the national laboratories or at the Goddard Space Flight Center or someplace like that, that still doesn't guarantee that university access will ensue, because the university access will be funded in research programs, basic research programs in the Department of Energy, probably not so much in NASA, but perhaps in Defense and the National Science Foundation and other places.

What happens is that funding to support those computer centers erodes away with time. You can mandate it today, but next year there will be many researchers submitting small research proposals, many more than can be funded, and these will start to drain money away from the support of the computer center to fund these $50 K or $100 K research efforts in universities.

Ms. **Lloyd**. In other words, we have a few problems we need to work out along this line before we advance?

Dr. **Cooper**. It is a major management problem.

Mr. **Trivelpiece**. Having had a minor role to play in establishment of the Fusion Energy Center in 1974 and 1975, we addressed the problem whether or not to fund the researchers and let them buy time on the computer or to make the fundible quantity time on the computer.

If you make it time on the computer, then the sort of evil that Dr. Cooper referred to can't occur, because it becomes the item that is given as a quantity to be managed and not dollars with the decision as to where the scientists could best spend them or where they would like to spend them.

Ms. **Lloyd**. Thank you very much for your excellent testimony and for answering our questions so well.

Mr. Walker?

Mr. **Walker**. Thank you. Just a question or priorities. As I sat here and listened, I hear a figure thrown around of 8 to 10 million dollars probably for supercomputers if we were to purchase them with the idea in mind of utilization along the lines that you are talking about.

Of course, we hear all the time about people who are looking for additional computer capacity in research and development work. Now, if we were to take the kind of money that would have to be invested in class VI computers and spread it out into a series of smaller computers, we can obviously buy a lot of capacity in that direction, too.

Which way are we going to get more R&D, buying into a series of smaller computers and satisfying all the various interests who are telling the committee they need those, or buying a series of class VI computers at considerable expense?

Mr. **Trivelpiece**. That is a clever question. I hope the answer I can provide is equally clever.

I think it has to do somewhat with the fact that the quality of research that you get would be different, the qualitative aspects of it would be different.

There are a collection of problems that are of national interest for which class VI or a subsequent generation of computer systems are the minimum buy in price in order to make realistic progress.

That is not to say, however, that if you took the same capability and spread it over a larger number of small problems, you wouldn't make progress in those problems, you clearly would. That progress would be valuable, too. There are a variety of ways to solve the problems, including using the class VI systems, but letting people use them for more menial tasks is generally not cost effective.

There are certain problems for which the class VI capability is needed and that is the only way you can make realistic progress, and you need a certain fraction of class VI machines and a certain fraction of the others, too.

It becomes a judgment as to which kind you need in order to meet the needs of the United States.

Dr. KERREBROCK. May I offer a slightly different answer? I think that the minicomputer and the microprocessor have become fully competitive in the marketplace. What you find is that unless you erect artifical barriers to the acquisition of these machines, the researcher will allocate moneys so as to acquire them.

One of our problems is that they want to do it faster than sometimes people think is advisable. You do not have to allocate special funding in order to insure the rapid utilization of commercial production small computers, because people will demand them.

On the other hand, the supercomputer has to have designated funding because there are very few users who can muster the amount of money that is required from their available budget to acquire a supercomputer.

So I think that they still require designated funding in order to make sure that we go ahead as fast as we should. Therefore, I do not think there needs to be a conscious prioritization between those two by the funding agencies.

Dr. COOPER. That is why DARPA has not funded supercomputer usage or applications for defense purposes, because there are not that many defense problems that require a class VI machine for their solution.

Instead, there are a wide variety of defense problems that require the kind of symbolic processing capability that we seek and it has only been recently that our problem appetites have outgrown the machines that are available.

Now, we can look beyond the lisp machines and the KL 10's that are currently available on the market to problems that we would like to solve that require thousands or tens of thousands of times the computer power that is available in these commercially available machines.

That is why we are launching out on this kind of supercomputer program, that is, the following of this line of intelligent machines that is represented by the lisp machine line.

Mr. WALKER. But if I understand correctly, based upon current needs we ought to be fairly suspicious if we have a whole series of people coming in telling us that they need supercomputer capacity.

You are saying that there are a fairly limited number of applications where supercomputer capacity is an absolute necessity and that we ought to be mindful of those, but that in many other instances that are needs for other types of computer capacity where we might better spend our money than on something in terms of a supercomputer so that we ought to be at least pretty discriminating as to where we put the money for supercomputers.

Is that an adequate summation of what I think I heard?

Mr. TRIVELPIECE. I am not sure where you are drawing the line quantitatively. You say a small number of problems. There certainly are, in some sense a large number of problems for which that type of supercomputer capability could be utilized. There are a finite number of people capable of addressing those problems, so at any given time in the United States you can't have an infinite number of supercomputers. That wouldn't be realistic. So it is to some extent matched to the number of people who are in a position to take advantage of and utilize them and I wouldn't characterize that as being a small number.

We can define what we mean by small, whether hundreds or thousands. I suspect there are probably thousands of people capable of properly utilizing them to address major large scale problems.

Mr. **Walker**. I wouldn't pretend to try to quantify the number either because I really don't know. The point I was getting to is that we tend to hear on the committee from research people who are interested always in being on the leading edge of technology and that is understandable, that is what we are all about here.

I guess what I am asking is should we be somewhat suspicious if we begin to get a whole series of people who are telling us that it is absolutely essential that they have a supercomputer in order to do what they tell us what they want to get done.

Mr. **Trivelpiece**. We are always suspicious when they bring us those proposals, too.

Dr. **Kerrebrock**. The market for the new generation of computer has been underestimated. It always turns out to be bigger than it is predicted to be.

Mr. **Walker**. It is pretty obvious that the market for the small microprocessor has been grossly underestimated from what we are now seeing happen, too.

Dr. **Cooper**. A point made in the Lax panel report is that there are opportunities to move supercomputation into new research areas where it is essentially unavailable today, and I think those opportunities could yield spectacular results for U.S. science. I think that what we need to do as research managers in the Government is to pick out those areas, to ferret them out, to recognize that they do exist and to make a good case to you that additional resources in that area are required and would have a potentially high payoff if we were to invest them.

I think that is true, and I would be happy some time to tell you about a number of opportunities that I think are there.

Mr. **Walker**. One final question that also follows the same line when we are talking about the investment being made. As we move into the kind of investment we would have to make per computer in each of these, I think we have a distinct question as to how quickly does that become significantly outmoded, and how quickly are we going to be moving into a class VII, class VIII, class IX that would tend to raise questions about our investment in class VI.

What do you see coming in the next several years that we ought to be thinking about as we talk about such investments?

Dr. **Kerrebrock**. Let me say from the viewpoint of our NAS activity we see a major change about every 3 to 4 years. We expect that, the computer we get in 1984 will have to be replaced in 1987 by something more powerful.

Mr. **Walker**. But that becomes a bigger and bigger problem as the computers become more and more expensive, too.

Dr. **Kerrebrock**. No, because they do not become more expensive. The canonical cost is about $25 million in 1984 dollars and you get a larger capacity in each generation.

Mr. **Walker**. Which, in fact, saves you money?

Dr. **Kerrebrock**. That is right. So it gets cheaper all the time in that sense.

Dr. **Cooper**. There is a wide variety of opinion.

Mr. **Walker**. We spend a lot of money around here to save money.

Dr. **Cooper**. There is a wide variety of opinion in the engineering community as to where these computers are likely to level off, and it is not clear that there is more than an order of magnitude increase in the computing speed that will be available with current technology.

It may require an expensive new departure of the kind that is inherent in the Denelcor design, which we heard about earlier today, but perhaps with some greater general purpose flexibility in it that will be required to go much beyond an order of magnitude.

So it may be that we will see a hesitation in the computer speeds in the next, say, 6 to 8 years at about the billion floating point operations per second level of computational speed, and it may require some inventiveness to get beyond that.

Mr. **Walker**. Thank you, Madam Chairwoman.

Ms. **Lloyd**. Thank you very much, Mr. Walker.

I want to thank the panel, again, and our previous panel, for their excellent participation this afternoon.

Without objection, we would like to ask if you would be amenable to answering questions in writing if some of our members would like to submit questions that were not here today.

We will continue our hearings tomorrow afternoon in this room at 2 o'clock, also. If there are no further comments, the subcommittee stands adjourned.

[Whereupon, the subcommittee recessed at 4:45 p.m., to reconvene the following day, Wednesday, June 15, 1983, at 2 p.m.]

COMPUTERS AND THEIR ROLE IN ENERGY RESEARCH—CURRENT STATUS AND FUTURE NEEDS

WEDNESDAY, JUNE 15, 1983

HOUSE OF REPRESENTATIVES, COMMITTEE ON SCIENCE AND TECHNOLOGY, SUBCOMMITTEE ON ENERGY DEVELOPMENT AND APPLICATIONS, AND SUBCOMMITTEE ON ENERGY RESEARCH AND PRODUCTION,

Washington, D.C.

The joint hearing met, pursuant to call, at 2:05 p.m., in room 2318, Rayburn House Office Building, Hon. Marilyn Lloyd (chairwoman of the Subcommittee on Research and Production) presiding.

Mrs. LLOYD. The subcommittee will come to order.

Good afternoon, ladies and gentlemen.

It appears to me that we are having these hearings to answer two major questions. The first is: How do we best use public and private sector resources to increase our Nation's computer capability in such a way as to essentially retain preeminence in the field?

As we heard yesterday and will hear again today, the stakes in terms of national well-being are truly enormous since the level of computer technology is a determinant in the rate of transition to a high technology society.

So many aspects of technology development are based on computer technology that this is an unavoidable fact.

The second question that we are asking is: How do we enhance access to our most advanced computers for public and private laboratories and the universities?

There is rather convincing testimony today that the accessibility to state-of-the-art computers at the universities is, in fact, a national disgrace, and we on this committee are well aware of the shortcomings in computer capability at the DOE National Laboratories and the NASA Centers.

We shall also hear more on that point this afternoon from the National Labs themselves.

These 2 days of excellent testimony have been most informative, but it seems to me that we don't have satisfactory answers to either of these major questions as yet.

I hope that this committee can work with the executive branch in bringing about, one, a focused and coordinated national effort to keep the United States at the forefront of computer technology that is a true "U.S. Incorporated" effort and; two, I hope that our two subcommittees can work with the Department of Energy in seeing to it that the most advanced computers are accessible to the national laboratories, universities and industrial contractors.

We are very sorry that the chairman could not be with us at this time but he certainly will be with us before the end of our hearings.

We do welcome all of you today and I welcome our first panel of witnesses and, without objection, your complete testimony will be made a part of the record. You can proceed to summarize or however you wish.

Our first panel includes Dr. John Roberts, Argonne National Laboratory; Dr. William Lester, from Lawarence Berkeley Laboratory, and Mr. Carl Haussmann, Associate Director, Lawrence Livermore National Laboratory.

We welcome you distinguished gentlemen and thank you for your participation this afternoon. Please proceed.

[The biographical sketch of Mr. Roberts follows:]

DR. JOHN J. ROBERTS

Dr. John J. Roberts is Associate Laboratory Director for Energy and Environmental Technology at Argonne National Laboratory. Under his purview are non-nuclear energy technology and related environmental control technology in the following six areas: magnetic fusion, electrical and chemical systems, fossil energy, advanced solar technology and biomass, solar heating and cooling applications, and energy conservation. He has been actively involved in R&D related to energy technology and environmental analysis since 1967. More recently his professional interests have focused on energy systems for communities. He has been an employee of Argonne National Laboratory since 1966 with the exception of three years (1971–74) during which time he organized and managed the Air Pollution Control Program for the State of Illinois. He received his undergraduate degree in electrical engineering from Rensselaer Polytechnic Institute and his Ph.D. in energy and applied math from the University of California in Berkeley.

STATEMENTS OF DR. JOHN ROBERTS, ARGONNE NATIONAL LABORATORY, ARGONNE, ILL.; WILLIAM A. LESTER, JR., PRINCIPAL INVESTIGATOR, LAWRENCE BERKELEY LABORATORY, BERKELEY, CALIF., AND A. CARL HAUSSMANN, ASSOCIATE DIRECTOR, LAWRENCE LIVERMORE NATIONAL LABORATORY, LIVERMORE, CALIF.

Dr. **ROBERTS**. Chairwoman Lloyd, I am pleased to appear before your subcommittee today at this hearing on the future of large-scale computing to present the viewpoint of Argonne National Laboratory and other multiprogram nonweapons national laboratories of the Department of Energy. Although I officially represent Argonne at this hearing, I have consulted with senior staff in the Computer Science Division at Oak Ridge National Laboratory in developing this testimony.

DOE Laboratories such as Argonne represent a unique class of research and development institutions wherein strong programs of fundamental research in the physical and biological sciences are coupled with nonclassified multidisciplinary applied science and engineering programs of high national priority. Thus, basic studies in physics, chemistry, biology, material science are pursued not only for their intrinsic value but also in conjunction with the development of fission, fusion, fossil, solar and conservation energy technology.

It is at this highly productive energy R&D frontier that we have examined the current and future opportunities for large-scale computing and the needs for major advances in the capabilities of computers.

We offer the following three recommendations:

One, that substantial supplemental funding be provided over a broad spectrum of DOE research and development programs to raise the level of computation to the current state of the art in mathematical science and computer hardware.

Second, that the Department of Energy should consider establishing a geographically distributed nationwide network of regional class VII computers and associated high-speed linkages sited at the multiprogram nonweapons laboratories and major universities.

And, third, that the Department of Energy should be supported in its current plans to substantially increase its program on long-term research and development in high-speed computing. That future generations of computers with novel hardware components and architectures should be procured at an embryonic stage in their commercial development and located at selected multiprogram national laboratories and major universities where specialized experimental software can be developed to exploit their advanced capabilities.

Although my written testimony provides substantial documentation and support for these three recommendations, I would like to present just some highlights at this time.

First, we have recommended that substantial incremental funding be provided to the scientists and engineers engaged in R&D across the full spectrum of DOE activities from energy research all the way to conservation. This call for support to the users echos remarks made by the panelists yesterday and by studies such as the Lax committee report.

The principal advantage which accrues to researchers with access to high-powered computers is the ability to simulate natural phenomena and engineered systems with increasingly realistic manner. Today, the researcher is faced with the practical limitations of those machines to which he can be assured reasonable access. He naturally scopes his problems to fit within the machine's constraints. He simplifies their dimensionality and in general tailors his research to the limits of the available machine rather than to the full potential of the underlying theory or to the state-of-the-art of computer technology.

From the prospective of nonweapons multiprogram laboratories, the staff at Argonne and Oak Ridge have summarized a number of representative areas of scientific and engineering research within which important advances await both enhanced access to the current generation of computers, supercomputers, and most importantly, the promised capabilities of the next decade.

Now, these summaries can be found in exhibit A of my testimony. I would like to mention one: Automated reasoning. Automated reasoning is one of the most sophisticated members of the class called nonnumeric computing. Most of the testimony that you have heard so far has focused on numerical computing.

In contrast to the number crunching aspects of most scientific and engineering analysis, nonnumeric computing entails the manipulation of words as in data banks; groups of words, as in language processors; formulas, as in algebra and calculus; and ideas, as in automated reasoning.

By reasonable extrapolation of our progress to date in developing computer software to perform automated reasoning tasks of problems as diverse as the proof of mathematical theorems and the efficient design of logic circuits, I foresee in 10 years full-scale control systems for facilities such as nuclear reactors, synthetic fuel plants, waste management facilities, which rely extensively on automated reasoning programs to digest, interpret, and direct the operators of those facilities to respond properly to the many simultaneous readings of sensors throughout the plants. I think that the probability of a core melt-down such as occurred at TMI would likely be reduced by over an order of magnitude when such controls based on automated reasoning are in place.

DOE and Argonne, in particular, are pioneers in this exciting branch of computer science.

A second recommendation concerns the establishment of a network of state-of-the-art supercomputers for unclassified DOE R&D, where such machines would be sited at multiprogram national laboratories such as Argonne and at major universities.

It is my personal opinion that such a network should serve the applied scientific and engineering programs of DOE as well as the energy research programs for two reasons. First, there is tremendous need for such computing power in the applied areas of DOE as well as in the energy research side. Second, and perhaps more importantly, a DOE-wide network would enhance the all important interactions between basic and applied researchers.

We recommend that this national network have a regional character as well, in contrast to siting the bulk of those machines at one or two laboratories.

This concept of a distributed network was echoed in the testimony yesterday of all three industrial representatives, agrees with the Lax report and is consistent with NASA's layout. We see the following advantages to a distributed DOE-wide network: close interactions between the computer systems staff and the users at the laboratory and surrounding universities; a wider experience in managing such machines and thereby providing broader future markets; and avoiding the migration of computer staff and users to a single region of the country.

Please keep in mind that such machines form an important part of the regional infrastructure supporting high technology enterprises. There are potential cost savings in high-speed data linkages within the region. Finally, there are at most minor incremental costs associated with integrating a supercomputer into an ongoing laboratory computer management organization such as exists at Argonne. That is, there is little return to scale by colocating all such supercomputers at one or two laboratories.

Our final recommendation addresses research in mathematics and computer science. I can summarize by observing that all the witnesses to appear thus far, and almost every recent study of this subject, urge the DOE to substantially enhance this area of research.

In particular I would like to reiterate the second part of our recommendation: that future generations of computers with novel hardware components and architectures be procured in an embryonic stage in their development and located at selected multiprogram national laboratories and major universities where specialized experimental software can be developed to exploit those advanced capabilities.

This concludes my prepared remarks. I would be pleased to answer any questions. Thank you.

[The prepared statement of Dr. Roberts follows:]

JOHN J. ROBERTS, ASSOCIATE LABORATORY DIRECTOR, ENERGY AND ENVIRONMENTAL TECHNOLOGY, ARGONNE NATIONAL LABORATORY

Chairman Fuqua, Ms. Lloyd, and members of the House Committee on Science and Technology. I am pleased to appear before you today at this hearing on the future of large-scale computing to present the viewpoint of Argonne National Laboratory and other multiprogram non-weapons national laboratories of the U.S. Department of Energy. Although I officially represent Argonne at this hearing, I have consulted with senior staff in the Computer Science Division at Oak Ridge National Laboratory in developing this testimony.

DOE Laboratories such as Argonne represent a unique class of research and development institutions wherein strong programs of fundamental research in the physical and biological sciences are coupled with non-classified multidisciplinary applied science and engineering programs of high national priority. Thus, basic studies in physics, chemistry, biology, mathematics, and materials science are pursued not only for their intrinsic value but also in conjunction with the development of fission, fusion, fossil, solar and conservation energy technology. It is at this highly productive energy R&D frontier that we have examined the current and future opportunities for large-scale computing and the needs for major advances in the capabilities of computers. Our principal conclusions and recommendations are as follows:

Conclusions:

1. Many areas of energy research and development of high national priority are being stifled through lack of adequate access to the current generation of supercomputers.
2. Outstanding opportunities exist for highly cost-effective applications of large-scale computing in the design and analysis of advanced energy technology.
3. Because of the intimate ties between hardware, system architecture, operating systems, and user software, the multiprogram non-weapons laboratories must become active participants in the development of future generations of computers.

Recommendations:

1. Substantial, supplemental funding should be provided over a broad spectrum of DOE research and development programs to raise the level of computation to the current state of the art in mathematical science and computer hardware.

2. The Department of Energy should consider establishing a geographically distributed nationwide network of regional Class VII computers and associated high-speed linkages sited at multiprogram non-weapons national laboratories and major universities to provide and operate a nationwide computing network for energy research and development.

3. The Department of Energy should be supported in its current plans to substantially increase its program on long-term research and development in high speed computing. Future generations of large-scale computers with novel hardware components and architectures should be procured at an embryonic stage in their commercial development and located at selected multiprogram national laboratories and major universities where specialized experimental software can be developed to exploit their advanced capabilities.

NEED FOR INCREASED SCIENTIFIC COMPUTING

RECOMMENDATION: Substantial, supplemental funding should be provided over a broad spectrum of DOE research and development programs to raise the level of computation to the current state of the art in mathematical science and computer hardware.

Our perspective as a national laboratory is that DOE should be highly motivated to support a national effort to promote the development of advanced computer hardware, system architectures and related software because of the many and significant scientific and engineering problems in energy research and development which call immediately for the state of the art in supercomputers and which will benefit greatly from improvements of two to three orders of magnitude in computing power by the 1990s. Further, for DOE sponsored research programs to contribute to the market for machines of the 1990s, they must begin now to greatly expand their access to the current generation of supercomputers and must become actively involved with industry in the evolution of the state of the art.

When the DOE Office of Energy Research recently advertised to its research groups at national laboratories and universities that it would make available substantial blocks of time in FY83 and FY84 on the Cray 1 operated by Livermore for the Magnetic Fusion Energy program, requests immediately exceeded the limited supply by a factor of ten. There is clearly a latent demand for access to supercomputers throughout the energy research and development community, and it is our opinion that the lack of adequate access, by which is meant sufficient time at acceptable cost, exerts a stifling influence on many areas of computational modeling and analysis. Further, in many areas of civilian energy research, the U.S. will fall behind the state of the art in scientific and engineering computations if this situation is prolonged.

Two points should be kept in mind: (1) the principal advantage which accrues to the researchers with access to high powered computers is the ability to simulate natural phenomena and engineered systems in an increasingly realistic manner; and (2) although large-scale computers may offer important increases in capacity (e.g., main memory), it is their promise of vastly superior speed, achievable by a combination of advanced hardware and novel architectures, which is most exciting. For example, although the IBM 3033s at Argonne have the capacity to handle more realistic modeling and analyses of the kinetics of fuel combustion, their capability (i.e., speed) is such that hundreds of hours would be required. Thus the researcher, faced with the practical limitations of those machines to which he can be assured reasonable access, naturally scopes his problems to fit within the machine's constraints--the dimensionality of the problem may be reduced, equations simplified, complex geometries approximated, and in general the research too often tailored to the machine rather than to the full potential of the underlying theory. The lack of adequate computer power causes researchers not to attempt at all the modeling of certain phenomena and processes. It is therefore important to the progress of energy research and development in the U.S. that researchers be provided substantially greater access to the current and emerging classes of supercomputers.

The Report of the Panel on Large-Scale Computing[1] and other recent assessments of this subject have documented many areas o research which will profit greatly from major advances in computing power. From the perspective of non-weapons, multiprogram national laboratories, staff at Argonne and Oak Ridge have summarized a number of representative areas of scientific and engineering research within which important advances await enhanced access to the current generation of supercomputers and, most importantly, the promised capabilities of the next decade. These summaries can be found in Exhibit A of this testimony and are highlighted below:

- Automated Reasoning--Research goals are threefold: understanding of the reasoning process; development of software to verify such understanding; application to problems such as design of logic circuits and diagnosis of complex physical systems such as nuclear power plants. These applications call for large memories and rapid execution of basic logical and numerical operations: realistic performance targets for the 1990s.

- Combustion Kinetics--Many of the important chemical species present in flames, including all of the principal chain carriers, are highly reactive and therefore difficult to study in the laboratory. Developments in theoretical and computational chemistry now provide a means of characterizing many of these reactions on the computer, thereby providing insights to assist in the design of laboratory experiments and increasing the accuracy of the computer models of combustion systems. The goal of computing the energetics of reactions to "chemical" accuracy (1 kcal/mole) requires an increase of two to three orders of magnitude in computing power.

- Critical Interactions in The Physics of Materials--The behavior of a material is the sum total of the behavior of the multitude of its constituent atoms, electrons and ions. Recent successes in solid-state and materials science have come about in large measure because the computer has increased our ability to account for both the complex forces and large numbers of particles involved in electronic and atomic motion in solids. While the computers of today allow a reasonable understanding of the gross features of simple pure systems in equilibrium, our ability to understand complex materials, materials with defects, liquids, and non-equilibrium phenomena depends critically on the availability of computers with higher speed and large memories. Adequate access to supercomputers is absolutely essential for the future viability of the U.S. programs in materials science and solid state physics.

---

[1] Peter D. Lax, *et al.*, "Report of the Panel on Large-Scale Computing in Science and Engineering" under the sponsorship of the Department of Defense and the National Science Foundation in cooperation with the Department of Energy and the National Aeronautics and Space Administration, December 26, 1982.

- High Energy Physics--Large-scale computing will provide techniques to investigate fundamental interactions of particles and fields with Monte Carlo methods and advanced symbolic programming systems. In particular, recent theoretical work on lattice gauge theories has benefited from, but is limited by, present fast computers.
- Protein Mapping Technology--Large-scale computing achieved by supercomputers or advanced, high-speed array processors are required to realize the potential of automated protein mapping technology for systematic discovery of disease markers and spinoffs to agriculture and genetic engineering. In general, biology is on the threshold of an era in which fruitful opportunities for computer-based analysis and requirements for computing power will both increase dramatically.
- Reactor Safety Research--The analysis of hypothetical core disruptive accidents, detailed three-dimensional analysis of fluid-structure interactions for calculating earthquake loads, three-dimensional fracture analysis of pressure vessels, and realtime simulation of reactor behavior are representative examples of computational studies of light water and breeder reactors wherein significant improvements in safety and significant reductions in plant construction costs and plant down-time may be achieved through the application of supercomputers.
- Magnetohydrodynamic Calculations for Magnetic Fusion--The design of large-scale fusion reactors requires a fundamental understanding of the time-dependent behavior of plasma instabilities. Advances in supercomputer technology will permit relaxation of approximations and simplifying assumptions in the solution of the coupled system of non-linear partial differential equations. The capability of handling models of greater sophistication could well lead to the cost-effective concentration of national research and development efforts on the most promising aspects of fusion energy technology. (These types of calculations profit greatly from advances in the state of the art of computing beyond that embodied in the Class VI computers at the National Magnetic Fusion Energy Computing Center at Livermore.)

For DOE and the nation to achieve the far-reaching benefits which would flow from enhancing the quality of computer-based scientific and engineering research, funding for such research must be significantly increased. Where appropriate, existing software should be adapted and new software developed to take advantage of the current generation of supercomputers.

Research budgets should support greatly increased access[2] to supercomputers.

Finally, we urge that such increases in funding to the "users" be *supplemental* to existing budgets and *not* be accompanied by a reduction in experimental or theoretical work which are essential to provide the data and methods necessary for meaningful computer analysis.

---

[2] The following section of this testimony recommends a nationwide supercomputer network.

NATIONWIDE SUPERCOMPUTER NETWORK

RECOMMENDATION: The Department of Energy should consider establishing a geographically distributed nationwide network of regional Class VII computers and associated high-speed linkages sited at multiprogram non-weapons national laboratories and major universities to provide and operate a nationwide computing network for energy research and development.

In order to realize the benefits outlined in the previous section, the nation's research scientists and engineers clearly need access to supercomputers. Researchers need to access supercomputers both to perform actual calculations and also to develop the computer models that will be used in the future. There are two important aspects to this development: The first involves determining how to represent and program the mathematical models of the phenomena being studied in ways that take best advantage of the "architecture" of supercomputers (i.e., they work internally in a somewhat different manner from conventional computers). The second is to determine exactly what features of an expanded model provide the most useful and valuable insights.

Most researchers do not have access to the currently commercially available (and proven) Class VI supercomputers such as the Control Data Cyber 205 or Cray Research Cray-1; although computer time is offered commercially, funds are not allocated to permit sufficient time to be available for the vast majority of calculations. In studying more advanced architectures, there are a number of groups supported to design and develop specialized supercomputers for their own project, or to do basic research into computer design; few researchers, however, have access to the systems of more advanced architecture which are becoming commercially available, such as the Denelcor Heterogeneous Element Processor and the Star Technology ST-100.

These needs are felt both at universities and the civilian national laboratories. We believe the needs can best be met by establishing a number of computing facilities equipped with supercomputers and by making them generally available to the research community. This is in accord with the recommendations contained in both the NSF's[3] "Prospectus for Computational

---

[3] William H. Press, *et al.*, report by the Subcommittee on Computational Facilities for Theoretical Research to the Advisory Committee for Physics, Division of Physics, National Science Foundation, March 15, 1981.

Physics" and the "Lax" report.* To disseminate the benefits of this new advanced technology as widely as possible, it is better to distribute these computers on a regional basis than to locate them all at one central location.

It will be valuable for users to have a center operating such a computer as geographically convenient to them as possible. It should be remembered that one of the functions of these centers is to enable people to learn how to exploit them. It is hard to identify and have available via a network all the information needed by the users of any new type of computer.

Finally, there is a need to provide experience in the operation of such facilities quite widely, so that the nation will have a cadre of people trained to operate supercomputers when they become less costly and unique. By operations we mean not only stopping and starting the computer, changing of magnetic tapes, etc., but also organizing the operating system software supplied by the manufacturer, integrating a supercomputer into operations with other computers, and allied activities.

To cover the various areas of the country would require perhaps six regional centers located in the major regional centers of research and development activity. DOE should determine the exact number and location, and the rate at which they should be installed. We believe that most major national laboratories are each technically equipped to receive and absorb a supercomputer relatively rapidly.

These supercomputers should be placed at existing installations, like multiprogram non-weapons national laboratories; DOE should avoid the additional difficulties and expense of setting up brand new facilities. Argonne would be a logical location for such a facility, and we are eager to install and operate a supercomputer as part of the recommended network.

These centers should have the most powerful production supercomputer available, and they should be accessible from any location. Interactive access would be from terminals via public carriers such as Tymnet and Telenet or via normal telephone calls. A high speed network using state-of-the-art communications technology would link the centers to major users and to each other. This network would have a high enough bandwidth for file transfer of all reasonable input and output, including graphics. Other users would access the regional facilities through general purpose university file transfer networks whose bandwidth limits the amount of input and output which can be sent; large files for users on these other networks, including, for example, any significant quantities of graphics, would have to be sent by mail. The smaller users would not generally be expected to need large files transferred very frequently.

The network would, in all possible cases, interconnect with existing campus and departmental facilities and networks (rather than with separate centers). In this way we avoid the costs of separate operations at each location to support access to the supercomputers. This would minimize the cost of making it available to the widest possible set of users.

---

* Lax, *et al.*, *op. cit.*

LONG-TERM RESEARCH AND DEVELOPMENT IN HIGH SPEED COMPUTING

RECOMMENDATION: The Department of Energy should be supported in its current plans to substantially increase its program on long-term research and development in high speed computing. Future generations of large-scale computers with novel hardware components and architectures should be procured at an embryonic stage in their commercial development and located at selected multiprogram national laboratories and major universities where specialized experimental software can be developed to exploit their advanced capabilities.

Effective use of the high speed computers of the 1990s and beyond will require long-term research efforts. Future generations of high performance computers will be quite different from the sequential and vector-oriented machines of the present; it will be correspondingly more difficult to utilize their full potential. Therefore it is essential that means be developed to (1) use them effectively, (2) improve future designs, and (3) distribute expertise on their use over as large a fraction of the country's technical community as possible. To achieve these goals DOE should support research on advanced computing and establish experimental computing facilities. The proposed major increase in FY 1985 of DOE support of these activities should not only be supported but supplemented by an increase in FY 1984 to permit a more timely and better staged entry into the field. This high speed computing initiative should be carefully coordinated with the existing DOE Applied Mathematical Sciences program. The latter supports activities that provide a sound foundation for an expanded program in advanced computing research.

The increasingly intimate ties between hardware, system architecture, operating systems, and user software create the need for user communities such as multiprogram national laboratories to play a major role in advanced computing. Furthermore, we believe that these laboratories will provide a fertile environment for research on high performance computers and dissemination of expertise on their use--a wide variety of computer applications, strong ties with universities, and increasing interactions with industry. The laboratories' experience in operating large user facilities such as particle accelerators will stand them in good stead in the advanced computing field as well.

Research on High Speed Computing Concepts

Computers with advanced architectures require substantial investment to attain their full potential. Class VI machines feature vector architectures; Cray-1, the most widely used, has been in use for over seven years, yet better algorithms and programming techniques are still being developed for them. The first Class VII machines will probably feature a few (2 to 4) parallel processors, each of which may be vector-oriented. Even this architecture is novel and there is no body of experience to guide its use. Future designs may have thousands of processors and memory units. There is general agreement that future architectures will be more complex and therefore more difficult to use initially--different languages, operating systems, algorithms, numerical methods, and mathematical models will all need to be devised. It will also be more difficult to determine the suitability of a particular design for a given class of computational models.

Research is needed to develop algorithms, techniques, and software that allow scientists to translate their models into programs without the necessity of learning the intricacies of these new computers and without the expense of a large staff of professional computer scientists. Such research, we believe, would best be carried out at a number of institutions, each bringing to bear its expertise in particular fields. For example. Argonne and Oak Ridge are well suited for such work as the result of their ongoing activities in applied mathematical analysis, computational mathematics and statistics, and automated program transformation.

At Argonne this research would include several activities. Work would continue on the investigation into algorithms for frequently occurring numeric and non-numeric computations that make possible the development of software that will execute reliably and efficiently on a variety of computing systems. This would represent a natural evolution of Argonne's work on mathematical software that yielded portable, reliable and efficient collections of programs that address a given area such as eigenvalue computations or solution of systems of linear equations. The value of these collections is indicated by their use in a wide variety of computer centers worldwide.

We would expand current research into techniques that facilitate the task of programming on advanced computers yet at the same time make it easier to move programs to different computing environments than it is at present. This research would have the explicit goal of not only moving application software to some currently available advanced architectures, but also of making it easy to move the software to future ones.

A starting point for this work might be the study of how to write very high level representations of algorithms and applications programs that can be (semi-)automatically adapted to various high-performance architectures. These adaptations to different architectures would be carried out for real problems. In doing this, one would explicitly have the goal of making the resulting program highly efficient on the chosen target architectures. The two goals, generality and efficiency, may sound incompatible, but preliminary investigations indicate that they can be achieved simultaneously. A concrete demonstration of their feasibility was provided by the efficiency of the Fortran implementation of a large symbolic manipulation program written in the

LISP program language. This pilot project was recently completed at Argonne using abstract programming techniques. At Oak Ridge research on the development of algorithms for new generations of computers would be the natural extension of their current work. Large-scale scientific problems typically have a structure that can be profitably exploited during the solution process. This fact has resulted in the emergence of sparse matrix technology as a new, important, active field of research. Oak Ridge's expertise in this field would be readily applied to develop algorithms for such problems utilizing parallel processors.

Oak Ridge is also engaged in developing innovative statistical methods and algorithms for existing classes of computers. Its research statisticians would turn their attention to developing new parallel computational techniques that may make it feasible to perform Monte Carlo simulations of complex processes, evaluate large computational models, and analyze massive amounts of data on a sound statistical basis.

Design Considerations

Design of the future generations of advanced architectures will be a formidable task. The complexity of these systems will require better automated design aids than are currently available. For the computer systems that are still decades away, we will even need to gain a better understanding of the fundamental physical processes involved in computation and information transfer.

Several design-related activities can be readily identified. By analyzing the shortcomings of current computer designs for certain problems, desirable new features can be invented and suggested to the developers of general-purpose advanced computers. For certain computing tasks, specialized architectures provide a cost-effective way to obtain high-speed computing facilities. The ability to move programs easily and efficiently to these new environments provided by techniques like abstract programming will make it more attractive to take advantage of such specialized architectures.

In some cases there are entire computations of interest to certain disciplines for which special purpose computers could be built that would be cost effective for that specific task. The scale of the computations required is such that, even on computers 50 to 100 times faster than Class VI machines, problems would run for hundreds of hours. The lattice gauge theory problems of theoretical physics are a prime example. Design of computers dedicated to these problems is underway at several institutions, including Argonne.

In other cases certain parts of a larger computing task occur frequently enough and consume sufficient resources to make it worthwhile to design specialized hardware to carry them out with performance approaching that of supercomputers at a small fraction of the cost. Today's attached array processors are the most widely known such machines and have found applications in several research programs at Argonne. This trend of building processors specialized for other functions will continue. An example of interest to Argonne is the implementation of hardware to carry out the time-consuming underlying logical and symbolic operations that occur in automated reasoning programs.

To date high-speed computers have been designed primarily for numerical computations. This is understandable since it has been those tasks that have demanded the greatest resources. As we enter a period in which non-numeric applications are becoming important, it is time to pay much more attention to the design of advanced computers for these tasks. Symbolic manipulation systems like MACSYMA have been found to be extremely useful in activities that involve complex mathematical expressions. These systems have the ability to perform the operations of algebra and calculus on much larger expressions than can be treated normally. So-called expert systems are being developed for applications such as medical diagnosis and organic synthesis. Automated reasoning techniques have been applied successfully to logic circuit design and validation. These non-numeric tools all require high execution speeds and large memories for certain applications. High performance computers for such tasks will have different architectures than the "number-crunchers."

One effort at Argonne might provide a source of expertise both on what kinds of advanced hardware are needed for automated reasoning applications (which can guide the design of future computers) and a source of problems to validate the success of the design. High speed in non-numeric operations alone is not the only consideration. For example, by 1990 it is possible that automated reasoning techniques would start to be applied to the control of large physical systems such as power plants or experimental apparatus. As currently envisioned this task might require not only fast logical inferences but also the rapid symbolic formulation and numeric solution of complex mathematical models to determine the state of the system. The resulting information would determine what line of reasoning needs to be carried out next by the reasoning component. Therefore, this application would create the need for computers that are extremely fast not only for logical inferences and other tasks associated with automated reasoning, but also in the execution of numeric computations, and for symbolic manipulation. In other words, we anticipate the need for "supercomputer" performance in one integrated computer system for numerical, symbol manipulation and reasoning tasks.

The task of designing advanced machines is becoming so complex that considerable automation is essential. This phenomenon creates the need for research on tools that will advance the rate of development of advanced processors. Among the many components of these tools are automated reasoning systems. Argonne would enhance its automated reasoning software and form relationships with groups developing user interfaces and knowledge acquisition components, and hardware development. This may make feasible the creation of powerful reasoning systems whose main uses would be tasks such as circuit design, systems design, program verification, and natural language processing--all tasks that bear directly on advancing the rate of development of the machines.

Finally, we see the need for theoretical studies to guide the design of high speed computer hardware in the decades to follow. For example, investigations in the fundamental physical limitations of the computation process may yield information on the extent to which the dissipation of energy into heat is a necessary part of the computation.

Experimental Computing Facilities

Experimental computing facilities are needed both to carry out the research and design activities described above and to provide opportunities and assistance in the use of advanced scientific computers to Laboratory staff, university faculty and students, and researchers from the private sector. Such facilities will operate experimental machines and be linked to existing production supercomputers; the latter will be the subject of study for some of the research. It may seem odd to experiment with a commercially produced computer to determine its behavior. Nevertheless, it frequently takes years after sophisticated machines are designed and built before their behavior is understood and they can be used effectively.

Each experimental computing facility would operate one or two leading-edge computers that promise to be particularly well suited to work of interest at the laboratory where they are sited. It would periodically acquire experimental high performance computers whose architecture promises to be effective for problems of interest to laboratory staff and their collaborators. These computers would be used to perform experiments designed to assess their virtues and vices. In some cases, hardware and firmware modifications would be required for the experimentation. Operating systems software and compilers would undergo frequent change. Close working relationships with the designers and manufacturers of the computer systems would be required for such experimentation. Proximity would be a distinct advantage to those who wished to use these experimental devices, at least in the early stages. The experimental facilities, coupled with the existence of groups doing research on algorithms, techniques, and software, would make it feasible to launch collaborative development efforts with manufacturers of advanced computing equipment, prospective users, and university researchers.

As soon as practical, these experimental machines would be made available to researchers at other sites through appropriate networks. An important feature of this activity is that members of the advanced computing research groups would collaborate with scientists and engineers in assessing the suitability of the experimental machines to eventually carry out production work. To facilitate this task, these experimental computing environments would be integrated with the local production computing facilities to the extent possible. This would reduce the significant difficulties commonly encountered in converting major applications programs to run in unfamiliar computing environments. The usual source preperation systems, output devices, libraries, and terminals could be used to run experiments on the new machines. For the projects whose goal is to devise algorithms and techniques that are suited to different classes of computers, it will be essential to have ready access both to the existing production supercomputers and the experimental ones. Another important benefit of such integration is that, if the experimental machines mature to the point that they (or similar ones) can be used for production work, it will be easier for the Laboratory's work to migrate from its current high performance computers to the new ones.

JJRjjp

EXHIBIT A

SUMMARIES OF REPRESENTATIVE RESEARCH PROBLEMS REQUIRING LARGE-SCALE COMPUTING

Automated Reasoning

Computation-Intensive Problems in Chemistry

Physics of Materials: Electronic Structure
Molecular Dynamics

Large-Scale Computing Needs of Theoretical High Energy Physics

Large-Scale Computing in Biology

Reactor Safety Research

Evaluating the Integrity of PWR Pressure Vessels

Time-Dependent MHD Calculations in Magnetic Fusion Energy R&D

AUTOMATED REASONING

P. Messina, ANL

STATEMENT OF RESEARCH GOALS

Automated reasoning is the study of concepts and techniques that enable a computer program to carry out tasks that have traditionally been thought of as requiring logical reasoning rather than numerical computation. There are three classes of specific research goals.

The first is to deepen our understanding of the components of the reasoning process. These have been broadly classified as representation of knowledge, rules for deducing new knowledge from existing knowledge, and strategies for focusing exploration and expansion of the knowledge base on the problem at hand. The second research goal is to develop automated reasoning software to take advantage of the latest in computer technology in order to verify that our understanding of the reasoning process can be implemented in computer programs that solve a continually expanding class of problems with acceptable speed. The third is to explore, and to help other scientists explore, applications of this technology in areas where deep reasoning is required. Such areas include the design of logic circuits, the verification of properties of computer programs, and the diagnosis of problems in complex physical systems such as nuclear power plants. In fact, automated reasoning is likely to become heavily used in the formulation, analysis, and validation of the design of future advanced computers.

LIMITATIONS OF CURRENT TECHNOLOGY

Automated reasoning technology has been successfully applied to difficult problems, including open questions in mathematics and logic. Many seemingly simple problems, however, require vast computer resources. sometimes strategies can be found that economize on such resources in the attack on a particular problem, but many problems require more resources than are available on today's machines. This is especially true of problems that require very prompt solutions such as real-time assistance with the operation of a nuclear plant. To analyze their behavior of systems with but a few dozen components requires seconds of CPU time on the fastest computers. It has been estimated that practical program verification systems must await two orders of magnitude increase in performance over today's levels. Computational resources needed for these applications are large memories and very great processing power (large number of basic operations performed per second).

NEED OF NEW GENERATION OF HIGH PERFORMANCE COMPUTERS

As in the numeric area, it is reasonable to expect that the increase in processing power required for advanced applications will be achieved primarily through multiprocessing. Automated reasoning algorithms are well suited to such an approach. Many tasks can profitably be done in parallel by a family of cooperating processes, much as a corporation can accomplish tasks too large and complex for a single person. The computers of the 1990s are expected to provide multiple processing units suitable for the execution of such algorithms.

Future generations of computers will also contain special purpose hardware. Just as floating point accelerators are routinely attached to today's computers, so we can expect special purpose hardware to be developed for basic, well-understood but complex operations that are performed millions of times by automated reasoning programs.

BENEFITS

The availability of much faster computers would expand the scale of problems to which automated reasoning techniques can be applied. This would benefit every application; for example, larger circuits, larger programs, and larger plant subsystems could be studied. A speed increase of three orders of magnitude, for example, may make feasible the real-time control of moderately complex physical systems. In many prospective applications the scale change is necessary in order for automated reasoning technology to be applied in an economically useful way.

Pure research in automated reasoning would also benefit. Ideas come from trying to solve hard problems. Expanding the class of problems that automated reasoning can successfully attack changes also the nature of the problems that are beyond its reach. These are the problems that spur theoretical progress. There is no prospect of running out of reasoning tasks; there are problems that will tax theoretical techniques and the fastest computers available for gerations to come.

COMPUTATION-INTENSIVE PROBLEMS IN CHEMISTRY

Thom. H. Dunning, Jr., ANL

Two of the programs in theoretical and computational chemistry at ANL which require substantial computing resources are:

(1) the kinetics of chemical reactions, especially reactions involved in the combustion of hydrogen and simple hydrocarbon fuels, and

(2) the physics and chemistry of transition metal clusters, especially metals relevant to catalysis.

The latter program is a joint program involving the Theoretical Chemistry Group, the Molecular Beam Group and the Surface Chemistry Group. The theoretical component of the program is to begin next fiscal year. A brief summary of each of these programs is given below.

KINETICS OF CHEMICAL REACTIONS

The development of accurate computer models of combustion processes requires the mechanisms and rates constants for the elementary chemical reactions involved over an extended range of temperature, from room temperature (300 K) to flame temperatures (>2000 K). Unfortunately, many of the species present in flames, including all of the important chain carriers, are highly reactive and are therefore difficult to study in the laboratory. On the other hand, developments in theoretical and computational chemistry begun in the late 1960s now provide a means of modeling many of these reactions on the computer. A strong theoretical and computational program in the kinetics of flames is all the more important because of the paucity of data available from experiment and because of the difficulty of interpreting much of the data available.

Over the past five years at ANL a number of important combustion reactions have been studied, including $O+H_2$, $OH+H_2$, $H+O_2$, $O+CH_4$, $H+CH_4$ (H_2+CH_3), $O+C_2H_2$, $H+C_2H_2$ C_2H+H_2), $H+H_2CO$, $H+HCN$ ($CN+H_2$), and $OH+CO$. These studies have primarily been concerned with the mechanisms and rates of the reactions, temperature and pressure effects, kinetic isotope effects, and the effects of vibrational excitation of the reagents. These studies have provided many new insights into the kinetics of these reactions; however, they have also emphasized the need for substantial increases in computer support as well as for further work in theoretical and computational chemistry.

For example, the above studies have shown that the goal of computing the energetics of reactions to "chemical" accuracy, $\sim$1 kcal/mole, remains as elusive as ever. Our studies indicate that present techniques with available computer resources (of several hundred hours of CDC 7600 time/year per theoretical group) are capable of an accuracy of a few kcal/mole for the types of reactions of interest here. While this often suffices for mechanistic considerations, and with semi-empirical corrections provides a basis for much of the needed quantitative rate constant data, the presence of such errors seriously detracts from the theoretical studies. Although studies directed toward the accurate calculation of bond energies and activation energies are few, certain general conclusions are suggested by those currently available:

(1) Multireference zero-order wavefunctions are required to consistently describe the reactants, transition state(s) and products of a chemical reaction.

(2) Many-body effects, which describe the correlations between the electrons in two constituents in a bi-molecular reaction, must be taken into account; changes in the intramolecular correlation energy are also important if chemical accuracy, i.e., ∿1 kcal/mole, is to be obtained.

(3) Large basis sets, or (at present unknown) improvements in the basis sets, will be required to properly describe the correlations of the electrons in chemical reactions.

The above three observations form the basis of our approach to computing the energetics of chemical reactions. The resulting calculations scale as approximately N^6, where N is the number of basis functions. Our experience suggests that basis sets two to three times larger than those in present use will be required to achieve chemical accuracy, with a concomitant increase of a factor of $10^2 - 10^3$ in computing time. For reactions such as those mentioned above, it is estimated that several hundred hours of CRAY-1 time may well be required to accurately compute the energetics of one of the more complex reactions.

Improvements in the description of the kinetics of these reactions is also a formidable undertaking. Beyond simple transition state theory models, the first step would be to carry out classical trajectory studies, thus explicitly including dynamical effects in the kinetics calculations. This introduces two problems: (1) determination of a global potential energy surface for the reaction, and (2) integration of the classical equations of motion, all in a hyperspace of dimension 3N-6 where N is the number of atoms. Aside from the difficulties noted above in the computation of potential energy surfaces, preliminary trajectory timings on a realistic surface for a typical five atom system indicate that such studies will require a substantial investment of computer time; again, a few hundred hours on a CRAY-1 computer. Finally, the classical description of the kinetics of the reactions ignores quantum effects, such as tunnelling, and the persistence of vibrational adiabaticity along the reaction path. These effects can also be expected to be important in the theoretical description of chemical reactions, at least for some quantities of interest. The appropriate quantum studies, using currently available techniques, would exceed the capacity of present computers, even supercomputers; new approaches are needed.

PHYSICS AND CHEMISTRY OF METAL ATOM CLUSTERS

In the late 1970s it became clear to a number of experimental and theoretical chemists that advances in laser and molecular beam technology and theoretical and computational chemistry would allow detailed studies of the physics and chemistry of small clusters of "naked" transition metal atoms. At present little is known of the physical and chemical properties of such clusters, even of such simple clusters as the dimers. Studies of the physical characteristics (spectra, energies, geometries, etc.) of, physi- and chemi-sorption on, and reactions involving these clusters would provide information on the molecular mechanisms of catalytic processes which simply could not be otherwise obtained.

The requirements for an accurate theoretical characterization of the structure of the clusters is not known; indeed, there is considerable argument at present over the various models which have been advanced to describe the electronic structure of transition metal clusters. The few calculations performed to date indicate, however, that it is indeed a challenging problem. To calculate the structure and spectra (electronic and vibrational) of the transition metal clusters multiconfiguration wavefunctions, and the inclusion of electron correlation effects, will be required. For the dimers such calculations are (just barely) possible, although the multitude of electronic states arising from the different occupations of the partially filled valence shells exacerbates the problem. Studies of larger clusters of transition metals will, however, require substantially more computer power than available at present.

SUPERCOMPUTING REQUIREMENTS FOR ELECTRONIC STRUCTURE CALCULATIONS IN MATERIALS RESEARCH

D. D. Koelling, ANL

The electronic structure of pure, perfect materials of modest complexity can now be calculated to good precision using existing computational machinery and algorithms. If one wishes to continue solving yesterday's problems, then sufficient access to current computing technology will be adequate. The problems which should be addressed are much more complex. But one would be very foolish to attempt such problems without access to a Class VI or better computer. A few examples can make it clear how this situation arises. One of the most useful tools the physicist has used in the study of electronic structure is the exploitation of the symmetry in the problem. However, the problems that are now of interest usually break that symmetry: defects, impurities, and surfaces all do so. Recently, Argonne has had greater involvement with defects and surfaces so I will choose them for my examples.

Defect studies have great relevance for radiation damage, metallurgical embrittlement and fatigue, and the thermal properties of materials. A most useful tool for studying defects is positron annihilation. Such experiments are ongoing at ANL and are currently being expanded in view of their utility. The positron annihilation data, however, overlays a number of different effects so computational predictions are invaluable as an aid to understanding the data. A landmark calculation for the defect structure in aluminum has recently been carried out at ANL. That calculation required a Herculean physical effort on the part of the researcher consisting of many all night sessions to get adequate access to computational resources. The truly interesting materials though are the transition metals and they require much more computational resources. The current realities are that that program can only be carried out for the refractory metals (niobium and molybdenum) granted an order of magnitude greater access to current computational machinery -- a shakey proposition at best. The extremely important metals iron and nickel are just not achievable in a finite time with current computational facilities: they need too much memory and computational power.

The second example is in the area of surfaces. The description of material surfaces is one of the most promising and rapidly growing areas of research: most of the interactions of matter occur at or near the surface. One of the most precise techniques for describing surface electronic structure has evolved at Northwestern University growing out of a bulk electronic structure package created at ANL. Certainly the Northwestern effort profited from the efforts of a group of bright hardworking researchers. But it is also profited from considerable access to Cray computers. The point can be clearly made by noting that an abortive attempt was made to bring the surface code back to ANL. It was abandoned because of inadequate computational power. We can run the bulk codes but not the surface codes. This is indeed unfortunate in light of the experimental interest at ANL in artificial structures. These systems are in fact multisurface problems and cannot locally be given full theoretical support.

The surface problem also illustrates another place where supercomputers are needed. In the close packed bulk calculations, the physicist knows enough about the solution to need only a minimal amount of variational freedom to get excellent results. This is not so especially in the surface layer. One needs to include great functional freedom and let the computation tell the researcher the nature of the problem. The whole history of surface states is evidence for this assertion. Including more basis functions increases the computational complexity for each iteration as the third power. In addition, one also faces an increased number of iterations as the search process explores the larger functional space. The result is that the computational requirements increase as the fourth or fifth power of the complexity introduced. This occurs in the defect, impurity, and even in the pure bulk calculations for more complex systems as well. In many cases, it is appropriate to pay the price. One needs to learn about the properties of defected solids, of surface layers, and even of the more complex solids -- ternaries as an example. It has always been my opinion that one of the tasks belonging to the national laboratories is to provide the resources necessary for such an effort. That cannot be done without much advanced computing power.

SUPERCOMPUTING REQUIREMENTS FOR MOLECULAR DYNAMICS CALCULATIONS IN MATERIALS RESEARCH

A. Rahman, ANL

A molecular dynamics study of a system of interacting particles as in a liquid or solid consists of conceptually placing a set of particles in a box and observing their motion. To do this, one must assume the nature of the interparticle interactions and then apply the equations of motion for very small time steps. The objective is to observe this model system and extrapolate its behaviour to macroscopic size. Compared to the few hundred particles which are the normal practical size studied, a macroscopic sample is an infinity. One can obviously greatly aid the extrapolation process by increasing the size of one's sample and in some cases this is essential. This is especially so when the particles act together in some fashion (such as forming a molecule) so that the sample of these entities is much smaller than the number of individual particles. Such an observation alone is adequate justification of the need for greatly increased computing power. But it is the emerging interest in more involved problems which really cries out for supercomputer capability. I offer two examples.

The structural transformations of solids critically involve long-range correlations of the motion of particles. An exciting breakthrough recently made at Argonne is the realization that the size and shape of the box should be included in the coordinates manipulated in the simulation stage of the molecular dynamics study. That improved technique has been proven out on small model systems at ANL and recently at the National Research Council of Canada (Solid He) and MIT (bcc iron). The time is ripe to apply the technique to metals and minerals at temperatures and pressures that are difficult to obtain in the laboratory yet of considerable import in our understanding of the natural world around us. Such studies are currently severely bottlenecked by limitations in computing power.

There is a rapidly developing thrust to understand the behaviour of isolated electrons in otherwise (classical) ionic materials. Work has begun on dilute metal-molten salt solutions to understand the solvation and transport of an electron in a molten salt system. Such information may have some use in the high temperature battery program. To do this problem, one introduces the new concept of doing a Feynman path integral for the quantum mechanical particle -- the electron. Although this can be done, it requires using many coordinates so it is as expensive computationally as adding 500 classical particles. To really study this problem however, it will be necessary to have more than one electron to study their interactions in the system. As such, it will become a problem far larger than any we have ever attempted before.

LARGE-SCALE COMPUTING NEEDS OF THEORETICAL HIGH ENERGY PHYSICS

E. May, ANL

Theoretical high energy physicists at Argonne are currently using Monte Carlo techniques to study solutions to quantum field theories. These theories are very difficult to study with traditional mathematical methods, and to solve the equations and get out reliable answers has required the development of new techniques. In particular, the use of these techniques allows the systematic study of non-perturbative effects, hitherto studied through approximations which often turned out to be invalid. The "lattice cutoff" scheme (first introduced by the recent winner of the Nobel Prize for Physics, K. Wilson of Cornell University) allows the calculation of Feynman integrals over all space-time to be replaced by a set or "lattice" of discrete space-time points. This use of large-scale computing has only recently been developed, and has already turned out to be highly productive and given considerable insight to these theories. The currently available computing capacity limits the size of lattices which can be studied, however. The application of a future large-capacity computing facility would remove the lattice size constraint and allow a determination of additional interesting effects in the theory. More importantly, it will allow new types of problems to be attacked. For example, the placing of scalar particle states on a lattice has recently provided many interesting results on the spectrum of bound gluon states; the availability of larger capacity computers would allow similar studies with fermions.

LARGE-SCALE COMPUTING IN BIOLOGY

L. Anderson, ANL

MOLECULAR BIOLOGY

Several major problems in molecular biology and biotechnology will require large-scale use of computing. The development of technology for mapping, classifying, and detecting mutations in the thousands of proteins which constitute the working parts of human cells entails large requirements for image processing and database manipulation. Work in this area at Argonne currently uses VAX and PDP 11 computers with attached high-speed array processors to achieve useful computational speed. Such a configuration is adequate for the completion of current feasibility studies, but will not support the full-scale project now being planned. Options for enhancing computer power to the required level include use of more powerful array processors, more array processors, or use of a supercomputer. The capability to maintain user interactiveness in image processing and database manipulation is a major goal. Long-term benefits to be derived from a comprehensive mapping project include the systematic discovery of improved disease markers, the development of complete lists of the protein working parts of human cells, leading ultimately to an understanding of cell function (only < 5% of these parts are presently known), and spinoff use of the protein mapping technology in agriculture (plant breeding) and microbiology (genetic engineering).

Also important will be rapid manipulation of DNA sequence databases of very large size (< 1 billion bytes). Such capabilities will be required both for research into the basis of gene expression and cancer, and for the development of genetic engineering approaches to the treatment of disease. In the very long term, substantial computational resources will be required to predict protein structures from sequence, thus initiating a new level of "protein engineering" capable of supporting the design of new enzymes and biological materials. The methods of protein structure prediction remain to be developed, requiring as they do an understanding of protein folding (a complex many-body problem). Nevertheless, the computing required will, by analogy with other similar problems, be substantial.

MOLECULAR COMPUTERS

The progressive reduction in the size of electronic components over the last decades has given rise to the view that computers may ultimately function using circuit elements of molecular dimensions. While the field of "molecular electronics" is presently in an early embryonic stage, the potential importance of such a technology in the construction of very fast and highly parallel supercomputers suggests that basic research in a national laboratory environment (combining chemical, biological and computing expertise) may be justified. Such research is likely to represent a risk level too high, and a time frame too long, to attract sufficient industrial interest.

THE TREND

In summary, biology is now on the threshold of an era in which computational requirements will drastically increase. This development is to be expected on account of the great complexity of biological systems as seen at the molecular level. As the variety of known molecular machines and informational elements increases toward the limit represented by a complete Human Protein Index and complete genomic DNA sequence, the effort to use this knowledge for the development of powerful new biological theories will increase. The scale of computing involved in such efforts is likely, by analogy with the somewhat less-complex problems of other sorts of engineering, to constitute a substantial fraction of national R&D computing requirements.

REACTOR SAFETY RESEARCH NEEDS FOR SUPERCOMPUTERS

F. Dunn, ANL

Three of the problems in reactor research at ANL that require the use of supercomputers are the analysis of HCDAs (hypothetical core disruptive accidents), the detailed three-dimensional analysis of fluid-structure interactions for calculating earthquake loads, and real time simulation of reactor behavior. In all of these areas significant improvements in safety and significant reductions in plant construction costs and plant down-time can be achieved by the use of supercomputers.

Currently, computer speed and memory size limitations greatly restrict the analysis in all of these areas. The computer codes currently used for HCDA analysis use one dimensional models of a single pin to represent all of the pins in a subassembly. Also, because of computer speed and memory limitations, about 15 subassemblies or groups of subassemblies are often used to represent the 200-500 subassemblies in a whole core. This modelling neglects pin-to-pin incoherence within a subassembly that would tend to mitigate the severity of an accident. Also, only part of the subassembly-to-subassembly incoherence is accounted for, and this incoherence also tends to mitigate the severity of an accident. In addition, because the results of multiple-pin in-pile safety tests only partly agree with the single pin model predictions, excessively conservative assumptions must often be made in whole core calculations in an attempt to compensate for model uncertainties. Also because of computer speed limitations, the three-dimensional calculations of earth-quake loads used to design reactor vessels and components are currently limited to very coarse models using a relatively small number of elements. This means that loads can not be computed accurately, and steel thickness must be increased to allow for uncertainties in the analysis. The increased steel thickness adds to construction costs and it increases thermal stresses. Again, because of computer speed limitations, current real time simulations of both water reactors and LMFBRs are limited to simple models that can not represent severe off-normal situations.

The approaches to using supercomputers in these areas involve using more detailed models in the HCDA and structural analysis areas, and obtaining real-time computing speeds for off-normal conditions in simulators. With the use of supercomputers it would be possible to develop multiple pin models for HCDA analysis to represent the incoherence within subassemblies. Application of these models to multiple pin safety tests should result in verified models that can be applied to whole core analysis without large uncertainties. A detailed three-dimensional analysis of the earhquake loads on a reactor vessel and its internal structures would show where the weakest points are. The weakest points could be re-designed or strengthened without adding excessive amounts of steel to other areas that are already adequate. Real-time simulations of severe off-normal reactor conditions can be used to test the adequacy of plant design, operating procedures, and operator training.

Since the cost of a supercomputer is only \$4 million - \$10 million, the benefits of using supercomputers can easily outweigh the cost. A detailed three-dimensional analysis of earthquake loads could lead to a vessel design that is safer, that is cheaper to build, and that has lower thermal stresses. A 10% reduction in the amount of steel in the vessel, vessel head, vessel support, and internal components could save enough in material and fabrication costs to pay for a supercomputer; and larger reductions should be possible. Similarly, a one month delay in the operation of a reactor due to licensing delays caused by uncertainties in safety analysis, or a one month shut-down due to inadequate design, faulty operating procedures, or inadequate operator training for severe off-normal transients could cost enough to pay for a supercomputer. More detailed and accurate HCDA analysis should reduce the uncertainties and excessive conservatism in current safety assessments and result in less costly containment designs.

EVALUATING THE INTEGRITY OF PWR PRESSURE VESSELS DURING OVERCOOLING ACCIDENTS

R. Bass, ORNL

APPLICATION DESCRIPTION

Under normal operating conditions, the reactor pressure vessel in a pressurized water reactor (PWR) is not exposed to temperatures and pressures likely to lead to the propagation of defects that might exist in the wall of the vessel. There is, however, a class of postulated accidents, referred to as overcooling accidents, that could allow cool water to come in contact with the inner surface of the vessel wall, resulting in high thermal stresses and a reduction in fracture toughness near the inner surface. This introduces the possibility of propagation of preexistent inner-surface flaws. This could lead in a worst-case situation to a rupture in the vessel wall and a possible meltdown of the reactor core. This possibility increases with reactor operating time because of the additional reduction in fracture toughness that results from exposure of the vessel material to fast neutrons.

A state-of-the-art fracture mechanics model has been developed to conduct parametric analyses of this condition and to evaluate several recorded PWR transients. Results of the latter evaluations have suggested that there might be vessels that have a potential for failure in a few years if subjected to certain kinds of transients. There is recent evidence to suggest, however, that this model might be excessively conservative and that the potential for defect propagation might not be as great as this model indicates.

CURRENT LIMITATIONS

The model presently in use allows some degree of analysis based upon a three-dimensional approach but requires a number of rationalized assumptions and estimates. A more definitive three-dimensional analysis would require the capability to include a much larger number of equation variables, allow assumptions of nonlinearity, and allow multistep calculations. But the development of such a model (involving complex interreactions among multidimensional, multieffect variables) requires access to the most advanced large-scale scientific computers available. Indeed, a very high level of model reliability wherein realistic simulations can confirm or disprove theoretical concepts requires access to computing resources which are now only in the experimental stage of development.

CONSEQUENCES

The consequences of the inability to perform the kind of analyses necessary to provide definitive answers has a number of potential implications with respect to our national interests. For this particular problem area, there are significant implications for public safety and cost. The ability, for example, to prove that a PWR has a useful life beyond what current models indicate could have extremely significant implications since the loss of a reactor through premature shutdown has been estimated to cost up to $1 million per day.

TIME-DEPENDENT MHD CALCULATIONS IN MAGNETIC FUSION ENERGY RESEARCH AND DEVELOPMENT

R. Hicks, ORNL

APPLICATION DESCRIPTION

The design of large-scale fusion reactors requires a fundamental understanding of plasma physics, particularly the time-dependent linear and nonlinear behavior of plasma instabilities. One technique for modeling plasma instabilities involves the use of time-dependent magnetohydrodynamic (MHD) calculations. These calculations can simulate, through the solution of a coupled system of eight nonlinear partial differential equations, the attainment of plasma equilibrium and stability for short time intervals (nanoseconds and microseconds). These calculations can indicate whether or not a particular MHD mode is stable, and if not its likely growth rate, structure, and potential destructiveness.

CURRENT LIMITATIONS

To make these equations manageable on existing computer equipment requires approximations and simplifying assumptions, including linearization, a reduction in dimensionality, and restriction to a particular geometry, ordering, or regime. If these simplifying assumptions did not have to be made, researchers might materially increase their understanding of plasma behavior and possibly reduce the amount of time required to develop fusion energy as a viable energy alternative. While much progress has been made in the development of improved numerical methods and computer codes as a way of reducing the requirement for approximations in these calculations, the key to more rapid progress is improved computer capabilities.

CONSEQUENCES

The availability of improved computing capabilities would allow the development and analysis of much more sophisticated models. This could lead to the conception of research and development efforts on the most promising forms of fusion energy technology and perhaps avoid the significant costs associated with the exploration of what might turn out to be less promising forms of technology. Moreover, there is a national interest involved in the timely development of an effective and economical alternative energy source.

Mrs. LLOYD. Thank you very much, Dr. Roberts, and also for your excellent paper.

We would like to hear from all of you and then we will have questions at that time.

Also, without objection, I would like to include the opening statements of our chairman, Mr. Fuqua, and also Mr. Winn and Mr. Sensenbrenner in the body of the contents of this hearing.

[The statements of Mr. Fuqua, Mr. Winn, and Mr. Sensenbrenner follow:]

STATEMENT OF HON. DON FUQUA

Today we continue our examination of the role and availability of state-of-the-art computers in advanced energy research, and the proper relationship between government and industry needed to promote the development of the next generation of supercomputers.

Yesterday we heard testimony from representatives of the computer manufacturing industry concerning the challenges facing continued U.S. leadership in this vitally important field. We also heard from government witnesses in agencies with research and development requirements that have had a major impact in the ongoing development of these machines, and where future R&D efforts will depend heavily on continued improvements.

Today we will hear from representatives of the DOE multiprogram national laboratories and the university community, where successful future research efforts will likely require expanded access to these machines. We look forward to your testimony.

OPENING STATEMENT BY THE
HONORABLE LARRY WINN, JR.

EDA/ERP JOINT HEARING ON SUPERCOMPUTERS AND THEIR ROLE IN ENERGY RESEARCH

JUNE 15, 1983

THANK YOU, MR. CHAIRMAN, FOR THE OPPORTUNITY TO WELCOME OUR WITNESSES TODAY. THEY COME PRIMARILY FROM THE NATIONAL LABORATORIES AND THE UNIVERSITY ENVIRONMENT. THIS IS IN COMPARISON TO OUR WITNESSES YESTERDAY WHO REPRESENTED INDUSTRY AND SEVERAL GOVERNMENT AGENCIES.

WITH A SMALL EXCEPTION, OUR WITNESSES TODAY SEEM TO AGREE THAT THE WAY TO MAINTAIN A PRE-EMINENT POSITION IN SUPERCOMPUTERS IS TO MAKE THEM ACCESSIBLE TO THE UNIVERSITY ENVIRONMENT. IT WOULD APPEAR TO ME THAT SUCH AN UNDERTAKING COULD INVOLVE SEVERAL UNIVERSITIES, POSSIBLY WITH THE ASSISTANCE OF A NATIONAL LABORATORY, TO ESTABLISH AND MANAGE A MULTIPLE ACCESS FACILITY FOR THE USE OF ALL. POSSIBLY WITH THIS AMOUNT OF PARTICIPATION THE EXPENSE TO NO ONE INSTITUTION WOULD BE EXCESSIVE. IN FACT, THE COMPUTER CAPABILITY MIGHT BE MORE COST EFFECTIVE THAN THE USE OF A SMALLER COMPUTER.

YESTERDAY TWO OF OUR THREE SUPERCOMPUTER MANUFACTURERS ADVISED THE COMMITTEE THAT THEY HAVE THE ABILITY TO DESIGN LARGER AND MORE EFFECTIVE COMPUTERS. THEY CLAIM THAT THEIR ONLY NEED WAS GOALS THAT SHOULD BE SET BY THE USER COMMUNITY. WHEN QUESTIONED, ONE OF THE MANUFACTURERS POINTED OUT THAT THESE GOALS SHOULD BE

SET WITH THE MANUFACTURERS SINCE THEY WOULD BE INTIMATELY INVOLVED IN THE PROGRAM.

THE THREE GOVERNMENT AGENCIES, DARPA, DOE, AND NASA, REPRESENTED ALL THE LARGE USERS OF DATA PROCESSING. THEIR NEEDS EXTENDED FROM NUMERICAL PROCESSING, OR "NUMBER CRUNCHING", INTO THE FUTURE OF "ARTIFICIAL INTELLIGENCE". THEIR NEEDS WERE SPECIFIED AND THE EXTENT TO WHICH THEY ARE PRESENTLY INVOLVED IN COMPUTER RESEARCH AND DEVELOPMENT WAS PROVIDED.

MR. CHAIRMAN, WE HOPE THAT TODAY'S WITNESSES WILL BE EQUALLY ENLIGHTENING AND, WE DO BELIEVE SO. WE WELCOME THEM ALL AND LOOK FORWARD TO THEIR TESTIMONY.

THANK YOU, MR. CHAIRMAN.

STATEMENT OF THE
HON. F. JAMES SENSENBRENNER, JR. (R-WIS.)

EDA/ERP JOINT HEARING ON SUPERCOMPUTER DEVELOPMENT

JUNE 15, 1983

THANK YOU FOR THIS OPPORTUNITY TO EXTEND A WELCOME TO OUR DISTINGUISHED PANEL MEMBERS.

WE LOOK FORWARD TO YOUR TESTIMONY ON THIS SECOND DAY OF HEARINGS ON THE PRESENT AND FUTURE ROLE OF COMPUTERS IN ENERGY RESEARCH.

ONE OF THE GREATEST TESTAMENTS TO THE SCIENTIFIC AND INVENTIVE CAPABILITY OF MAN HAS BEEN THE DEVELOPMENT OF THE COMPUTER. ALTHOUGH THE COMPUTER IS NOT A NEW TECHNOLOGY, WE ARE ONLY BEGINNING TODAY TO FULLY DEVELOP THE POTENTIAL OF THE COMPUTER. THROUGH A PARTNERSHIP OF ACADEMIA, INDUSTRY AND GOVERNMENT, THIS COUNTRY HAS ACHIEVED A PREEMINENCE IN THE COMPUTER FIELD. HOWEVER, OUR PREEMINENCE IN THIS FIELD IS BEING THREATENED.

I AM PARTICULARLY CONCERNED OVER REPORTS THAT THE COMPUTER INDUSTRY IN THIS COUNTRY IS UNABLE, EITHER BECAUSE OF A LACK OF FINANCIAL CAPITAL, OVERLY RESTRICTIVE ANTI-TRUST LAWS AT HOME, OR COMPETITION FROM OTHER COUNTRIES, TO CONDUCT THE BASIC RESEARCH THAT IS REQUIRED FOR THE DEVELOPMENT OF THE NEXT GENERATION OF COMPUTERS. THESE REPORTS, IF ACCURATE, WILL SERIOUSLY JEOPARDIZE BOTH THE ECONOMIC AND TECHNOLOGICAL FUTURE OF THIS COUNTRY.

I LOOK FORWARD TO YOUR THOUGHTS ON THIS SUBJECT.

Mrs. **Lloyd**. Dr. Lester, we look forward to hearing from you now.

[The biographical sketch of Mr. Lester follows:]

Biographical Sketch of William Alexander Lester, Jr.

William A. Lester, Jr., is a professor of chemistry at the University California, Berkeley, and a faculty senior scientist in the materials and molecular research division of Lawrence Berkeley Laboratory.

A native of Chicago, Lester received his early education in the Chicago Public Schools System. In 1958 he was graduated from the University of Chicago with a bachelor of science degree and the following year earned a master of science degree there. He then studied at The Catholic University of America, earning a doctor of philosophy degree in chemistry in 1964.

Lester's professional experience began as a researcher in quantum chemistry for the National Bureau of Standards while pursuing graduate studies at Catholic University. From 1964 through 1968 he was associated with the University of Wisconsin—doing research in molecular collision theory, molecular quantum mechanics and scattering theory, teaching undergraduate physical chemistry and graduate courses in theoretical chemistry and serving as assistant director of the Theoretical Chemistry Institute.

In 1968 he joined the professional staff of IBM Research Laboratory in San Jose, California, doing research in quantum chemistry and molecular collisions. In 1975 Lester moved to the IBM T.J. Watson Research Center in Yorktown Heights, N.Y., where he was on the technical planning staff of the director of research with responsibility for science in all IBM research division laboratories (Yorktown, San Jose, and Zurich). From 1976 to 1978 he was manager of the molecular interactions group in the San Jose Lab.

Lester, in 1978, was chosen as director of the National Resource for Computation in Chemistry (NRCC), located at the University of California's Lawrence Berkeley Laboratory. In that position he organized and directed the nation's first unified effort in chemistry, concentrating on improving efforts of chemists by making available information on new computational methodologies, development of computational algorithms and related sofeware, as well as conducting research in chemistry that utilizes computational methods. He also held the position of associate director and staff senior scientist at LBL.

He is a member of the recently formed Task Force on Large Scale Computing of the American Chemical Society's Joint Board-Council on Science, and serves on the National Research Council's Committee to Survey Opportunities in the Chemical Sciences with responsibility for computers in chemistry.

He is a member of the American Chemical Society, the American Physical Society, the National Organization for the Professional Advancement of Black Chemists and Chemical Engineers, Sigma Xi, National Institute of Science, and the American Association for the Advancement of Science.

This fall he will receive the 1983 Alumni Achievement Award in Science of the Catholic University of America Alumni Association.

In 1979 he was winner of the Percy L. Julian Award in Pure or Applied Research, presented by the National Organization for the Professional Advancement of Black Chemists and Chemical Engineers.

His many professional activities include the following: Specialist Editor, Computer Physics Communications, 1981; Chemistry Division Advisory Panel, National Science Foundation 1980–81; United States National Committee, International Union of Pure and Applied Chemistry, 1976–79; Chairman, Division of Physical Chemistry, American Chemical Society, 1979; Chairman, Gordon Conference on Atomic and Molecular Interactions, 1978.

Lester lives in Oakland, California.

Dr. **Lester**. Thank you.

Madam Chairman, and honorable members of the committee:

It is a pleasure for me to be here today to testify on the subject of computers and their role in energy research: current status and future needs.

In the time allotted me for oral testimony, I would like to touch on some of the highlights and offer some personal comments on my written testimony. And my comments are couched in the order of the questions that were developed. Let me read those and comment briefly.

The first is: To what extent and in what quantities should state-of-the-art supercomputers be made widely available to energy science and engineering researchers at national laboratories and universities?

My remarks are based upon consultation with Lawrence Berkeley staff as well as staff at Brookhaven National Laboratory.

There are two reasons for making supercomputers available. First and foremost, is the need to do computations to get research done.

Second, is the need to train students and potential researchers for future unspecified tasks.

Why are they not available now?

Basically it is a combination of money and bureaucracy.

Scientific computers should not be treated as special objects. They are research tools and their costs, are part of the cost of doing research are part of the cost of training and education. As such, they should simply be budgeted appropriately but not subject to any special scrutiny. With modern networking, for most problems the physical location of the computer is not critical.

The question of whether the operation should be a commercial one or government-run should primarily be one of cost effectiveness. It should be noted, however, that since we are discussing state-of-the-art supercomputers, the high risk associated with such frontier activities is likely to rise the commercial costs to an unreasonable level, so that there may be no commercial option.

To respond to the second question and here I refer to the question: What role can the Department of Energy play as a buyer and/or developer of technology for the next generation of computers?

From my extended response in the testimony it should be clear that the Department of Energy can have significant impact on the development of the next generation of supercomputers as both a buyer and a developer.

It is documented in the Lax report and elsewhere that there are research problems in universities and in national laboratories that require the capability of present supercomputers and others that need many times present supercomputer performance.

As a buyer of supercomputers, the Department of Energy can effectively reduce the cost by absorbing some of the risks which otherwise inhibit manufacturers and by providing a market for unproved products.

Equipment purchase would provide the tools to the research community, bring the community into intimate dialog with the manufacturers on needs, requirements, and shortcomings of present equipment-invaluable information for product improvement. The expanded market made possible by the Department would facilitate increased investment in research and development of future products that may compete with foreign competition.

There is also a role for the Department of Energy as a developer of technology for the next generation of computers and I describe in my testimony two aspects at Lawrence Berkeley Laboratory.

One entitled Alternative Approaches to High-Speed Computing, and the other entitled Research in High-Speed Computers.

To move quickly along to the third and fourth questions which are: What are the current status and future prospects of federally and privately funded computer R&D? And, is the private sector's research directed toward development of computers that will be needed for forefront research being pursued in DOE programs?

In reply let me make the following statements. There is considerable activity and computer R&D in both federally supported work and in the private sector. Because of defense needs of the Federal Government and business requirements of the private sector, the strongest emphasis has been on information handling, communication, and automation.

There has been a long hiatus in the advancement of computational power for numerical problems. The efforts of Seymour Cray providing a notable exception.

Now there appears to be a growing wave of interest in computational tasks driven in part by computer aided design systems and the private sector shows more enthusiasm. There is urgent need for a corresponding renewal of federally supported R&D and the technologies for designing, building, and utilizing new generations of supercomputers .

The Department of Energy has the research missions which will need new generations of supercomputers and the research staff to pursue these developments.

That concludes my statement.

[The prepared statement of Dr. Lester follows:]

STATEMENT OF DR. WILLIAM A. LESTER, JR., PROFESSOR OF CHEMISTRY, UNIVERSITY OF CALIFORNIA, AND FACULTY SENIOR SCIENTIST, LAWRENCE BERKELEY LABORATORY

Mr. Chairman and honorable members of the committee, it is a pleasure for me to be here today to testify on the subject of "Computers and Their Role in Energy Research: Current Status and Future Needs".

In the time allotted me for oral testimony, I would like to touch on some of the highlights and offer some personal comments on my written testimony, which expresses my views on this subject in considerably more detail.

I will attempt to answer in order the questions posed by Chairman Fuqua and Chairman Lloyd in their invitation. As might be expected, my responses will be of greater length in those areas where I feel our expertise and experience can be most useful.

1. To what extent and in what quantities should state-of-the-art supercomputers be made widely available to energy science and engineering researchers at national laboratories and universities?

There are two reasons for "making supercomputers available." First and foremost is the need to do computations to get research done. Second is the need to train students and potential researchers for future unspecified tasks. Why are they not available now? Basically it is a combination of money and bureaucracy. Scientific computers should not be treated as special objects; they are research tools and their costs are part of the cost of doing research or part of the cost of training and education. As such they should simply be budgeted appropriately but not subject to any special scrutiny. With modern networking, for most problems the physical location of the computer is not critical. The question of whether the operation should be a commercial one or government-run should primarily be one of cost-effectiveness. It should be noted, however, that since we are discussing state-of-the-art supercomputers, the high risk associated with such frontier activities is likely to raise the commercial costs to an unreasonable level, so that there may be no commercial option.

It is important to note here that in the molecular sciences significant amounts of Class VI (Cray-1) time are being made available to researchers in the United Kingdom through the Science and Engineering

Research Council's Daresbury Laboratory. This laboratory is at the center of a network providing large scale computational facilities for universities. Similarly, the Institute of Molecular Sciences in Japan, serves as the center of a national computer network providing access to two fast (CPU speed of a CDC 7600) large memory computers. The IMS has already announced that it will receive one of the first Cray-1 class computers made in Japan within the next year or two.

To provide insight on the current status and future needs of energy science researchers, I summarize below, as an example of the requirements of a DOE multiprogram laboratory, the Lawrence Berkeley Laboratory's response on this issue. Comments are provided for research programs concerned with the modeling and mathematical simulation of complex physical systems found in solid state physics, quantum chemistry, geosciences, and applied mathematics.

Pseudo-Potential Studies

At the present time, effective or pseudo-potential applications for the solid state are limited to elemental systems. The availability of a very high speed computer would make possible the extension to compound semiconductors, to transition metals, and to more complex crystal types and would permit the study of systems having more atoms per unit cell. In addition, the computation of surface states becomes realizable, and with this the prospect of providing detailed information on catalysts of chemical interest. Such work would represent a considerable advance over computational studies of systems currently the subject of investigation in ultrahigh vacuum apparatus.

Monte Carlo Simulations

Many areas of chemistry and physics are being advanced by the use of Monte Carlo techniques which provide benchmarks against which to test approximations and new information in cases where experiments cannot obtain accurate data. Computations on a high speed computer can make possible studies of many-body systems in condensed phases (liquid, solid, metallic) and at interfaces. Simulation studies have played a crucial role in the elucidation of phase transitions and, with the availability of a Class VI or better machine, the complete fermion system (electrons and nuclei) becomes amenable to treatment, as evidenced in pilot studies by LBL scientists.

Extensions of time-dependent phenomena at the microscopic level, leading to rigorous calculations of electrical and thermal conductivities by means of response functions, become practical with a Class VI or faster computer. In addition, the complicated area of finite temperature many-body systems also becomes accessible. Simulations can significantly benefit from increased central processor unit (CPU) capability. Unlike many other areas, improved accuracy and reliability of Monte Carlo calculations are critically dependent on CPU cycles.

Monte Carlo methods are also beginning to be used for the study of electronic structure and interactions of atoms and molecules, utilizing quantum as well as semi-classical and classical theories. LBL researchers are pioneers in his effort. The CPU capability of at least a Class VI computer will allow a thorough testing of this method.

Ab Initio Electronic Structure Determination and Potential Energy Surfaces

Molecular theorists have been in the forefront of the use of powerful computers to determine the electronic structure and properties of atoms and molecules. New methods for describing electron correlations in large systems and in excited states, including the graphical unitary group aproach for reducing the difficulties associated with configuration interaction calculations and multi-configuration Hartree-Fock methods, are being spearheaded by LBL researchers. Furthermore, traditional methods are being applied to larger and more complex molecules and interacting systems with useful benefits. In addition, LBL scientists have been successful in developing new automated methods for finding reaction barriers, transition states, and equilibrium geometries of molecules.

These developments presage a period of exciting developments that can have significant practical impact in areas such as combustion and atmospheric chemistry when the full power (both memory and speed) of a Class VI system is made available

Collision Dynamics

Advances in quantum, semi-classical, and classical methods for determination of inelastic collision events have reached the stage where theory is an essential adjunct to experimental verification of models. Theory further serves as a source of new information not obtainable by experimental methods for systems for which accurate potential energy surfaces can be computed. For molecule-molecule energy transfer studies, progress will be governed by the computational power available.

Quantum mechanical reactive scattering calculations (requiring large memory storage as well as raw computing speed) present even more of a challenge to computational facilities. The methodology has progressed to the stage that, with Class VI or better facilities, reliable reactive cross sections for simple atom-diatom systems are possible. There have also been important advances in theoretical methodology for reactive processes in polyatomic molecular systems. These approaches interact closely with the quantum chemistry techniques for generating reaction paths (see previous section) and involve various dynamical approximations, but they are an important first step toward dealing with more complex systems.

Geosciences

Research in the field of earth sciences requiring the capability of a modern, Class VI or VII machine is presently concentrated in two main areas:

(i) numerical modeling of subsurface systems

(ii) processing and analysis of geophysical data.

Current research in these areas addresses ever more complex systems, emphasizing a greater degree of realism, completeness, and detail. As a consequence, existing computational software is applied to more difficult problems, while ongoing research develops new large computer programs with greater demand on computational hardware. Typically the most computationally intensive work involves the solution of very large sets of coupled equations. Computational tasks of a similar nature arise in petroleum engineering (exploration, field development). These kinds of calculations have provided a strong incentive for the development and application of computers with vector processing capabiltiy. The so-called "vector-machines" can simultaneously perform a large number of operations which arise in the solution of systems of equations. This capability can improve computational speed by a factor of 10 or more in comparison with scalar computers.

Applied Mathematics Research

Research programs in applied analysis, computation mathematics, and numerical methods for partial differential equations at LBL emphasize numerical studies and techniques applicable to computational fluid dynamics, in particular combustion modeling. The combustion studies are directed towards the construction of realistic models that are suitable for comparison with experiments and are capable of being used for design purposes. Large-scale scientific computing capabilities are clearly required for research of this type. Anticipated growth of these research programs will require continued and expanding use of large-scale computational resources.

The Cray-1 computer at the Lawrence Livermore National Laboratory is currently being utilized via ARPANET to develop numerical methods based on explicit local wave tracking for computing compressible fluid flows. Vector computers such as the Cray-1 have the capacity to perform realistic calculations because the numerical methods used have been chosen to maximize performance. Furthermore, as the physics being modeled becomes more complicated, the amount of computer time required to explore the range of design options for these methods will increase substantially. For the methods currently being investigated, one hour on the Cray-1 is approximately 10 hours on the LBL CDC 7600.

2. What role can the Department of Energy play as a buyer and/or developer of technology for the next generation of computers?

From my response to the previous question, it clearly follows that the Department of Energy can have significant impact on the development of the next generation of supercomputers as both a buyer and a developer. In universities and in national laboratories, it is documented in the Lax report and elsewhere that there are research problems that require the capability of present supercomputers and others that need many times present supercomputer performance. As a buyer of supercomputers the Department of Energy can effectively reduce the cost by absorbing some of the risks which otherwise inhibit manufacturers and by providing a market for unproved products. Equipment purchase would provide the tools to the research community, bring the community into intimate dialogue with the manufacturers on needs, requirements, and shortcomings of present equipment -- invaluable information for product improvement. The expanded market made possible by the Department would facilitate increased investment in research and development of future products that may compete with foreign competititon.

There is also a role for the Department of Energy as a developer of technology for the next generation of computers. Described below are research directions of two Berkeley research groups that can assist this role entitled:

(a) Alternative Approaches to High-Speed Computing, and

(b) Research in High-Speed Computers.

Alternative Approaches to High-Speed Computing

Most experts agree that the speed of standard or serial computers is approaching fundamental physical limitations imposed by signal propagation (the speed of light) and heat dissipation. This is evidenced by the fact that the fastest modern serial computers are only 2 to 3 times faster than the CDC 7600 introduced in 1968. Although it is possible to increase present serial processing speeds by a factor of 10 to 50, such increases will require new technological advances and will probably be relatively expensive.

The Cray-1 computer, introduced in 1976, attempted to circumvent these obstacles by performing various identical operations in parallel (vector processing). Suppose, for example, a problem required that two sets of 50 numbers be added, by pairs, and the 50 results were, respectively, to be multiplied by a third set of values. A vector machine would use 50 separate addition units and multiplication units and perform the entire operation in two steps rather than 100. For the portion of a problem that can be organized in such a manner, this technique yields substantial increases in speed over the traditional serial machine. If, however, it was necessary to examine the result of each addition, for example to determine the appropriate multiplication factor, the performance of the vector processor would diminish considerably. Software effort in both the United States and Japan is being directed to developing codes, called compilers, that attempt to reorganize standard programs in order to permit more effective utilization of vector machines.

The extent to which problems are amenable to this approach varies considerably and a great many important physical problems appear to be essentially non-vectorizable. The advances in micro-electronics during the past decade have made new approaches to computing feasible. This involves using multiple computers, or processing elements, collectively on the same problem. Unlike the vector approach, such processors are not constrained to perform the *same* operations at the same time and, may, for example, be working on different facets of a calculation simultaneously. Whatever speed individual serial computers ultimately achieve, it is clear that multiprocessor machines could potentially extend this speed by factors of hundreds or thousands.

Although a number of manufacturers now provide systems which incorporate, or can be expanded to include, several processors, such systems (with the exception of Denelcor's HEP Computer) are designed primarily to run separate problems and not to work collectively on the same program. The potential power of multiprocessing computers thus remains virtually unexplored within the commercial marketplace. This is primarily due to the difficulty in architecturally organizing many independent processors so that they may work productively towards the same goal. The main problem areas in this respect are in multiprocessor control and inter-processor communication.

Although many designs for such multiprocessor systems exist, few have actually been constructed. The reason for this is both the expense involved in constructing such a system and the fact that many proposed designs are directed at specific classes of problems and hence may not be commercially viable. The new industrial consortium, the Micro-electronic and Computer Corporation, currently involving about a dozen corporations, is vitally concerned with the lack of development in this area. A recent MCC report stated that "To date, attempts to improve performance through highly parallel structures has been relatively disappointing. We believe the major reason for this lack of progress is the high real and personal cost to build and evaluate parallel structures". This fact, coupled with newly expanding computer applications in both business and personal computing, make high-risk, high-priced ventures appear unnecessary.

The need for higher-speed, more cost-effective computing continues to grow in basic scientific research. The analysis of the experimental data collected at CERN (European Organization for Nuclear Research) which led to the recent breakthrough in high-energy physics, the discovery of the W particle, will require several years of processing time on their computers and is currently utilizing most of the computational capability of that facility on a continuing basis. Parenthetically, the analysis of this information does not lend itself to a vector processing approach, but would be quite amenable to a multi-processor architecture.

The need for higher-speed and more cost-effective computing has, in the absence of commercial alternatives, led some research organizations to develop their own solutions. Lawrence Berkeley Laboratory's Nuclear

Science Division, facing problems not unlike CERN's, began in FY80 to develop a multiprocessor system called MIDAS (Modular Interactive Data Analysis System). This system combines the central processing units from commercial computers with some independently developed, special purpose processors and a specially designed communications system. The architecture permits considerable flexibility in organizing the processing elements for different applications. Thus, the computing structure itself may adapt, within limits, to the requirements of the problem. For the purposes of control, MIDAS organizes groups of processors into clusters under the guidance of a standard commercial mini-computer. One such cluster, containing 10 processors, is currently operational. In tests utilizing existing FORTRAN programs the performance of the cluster was approximately equivalent to that of CDC 7600 (at about 1/20 the cost). A variety of programs ranging from a Monte Carlo fission simulation to an optimal path determination have also been tested. It appears that the system has applications in many scientific areas, such as quantum chemistry, statistical mechanics, theoretical physics, geophysics, transport theory, plasma physics, etc.

Perhaps the most exciting aspect of the system is its potential for expansion. The existing cluster may be expanded to contain additional standard processors (and faster ones as they become commercially available) as well as multiple vector processors. The architecture further permits the future addition of special hardware processors to perform particular mathematical algorithms that may prove difficult to break into parallel components. A cluster may thus potentially be expanded to perhaps 20 or more times its current capability. Since multiple clusters could be constructed and interconnected, the eventual capability is difficult to predict.

Industry appears quite interested in the MIDAS project and would like to work with LBL on its future development. Obviously this permits a relatively painless industrial assessment of the approach and its future commercial viability. This is an important distinction since the project could be a complete success from the perspective of scientific research (i.e., high speed and cost effective) and still be commercially impractical. Thus, such joint efforts are necessary if we wish to bridge the gap to the next generation of computing. In the Dec. 17, 1982 issue of Science, Buzbee, Ewald and Worlton, from LANL stated:"To maintain U.S. leadership, private and government-sponsored research is required to identify areas for development that are likely to yield technical and market success, thereby lowering the investment risk to a point where the private sector will take over". This sentiment is also echoed by industry itself: "The current trend with the U.S. computing industry is to improve our current products and technology incrementally. This approach will not result in the development of new computer systems and technology that will be required for the next generation of computing systems" (report of MCC's Alpha-Omega Project, Dec. 1982).

Research in High Speed Computers

1. Introduction.

High performance computing research in the United States is at a crisis stage. Reearch and development, especially in universities, has declined rapidly over the last few years. This decline, if allowed to continue, could produce two very unwelcome results: (1) foreign competitors could easily become the technological leaders in high speed computing, and (2) problems requiring the use of high speed systems will not be solved.

We believe that there are three necessary components of a potentially successful high speed computer research program. (1) There is a need for support for problem-centered architecture design; (2) there is a need for support for building physical prototypes; (3) there is a need for the support of experimental computing facilities. These components are discussed below.

2. Problem-Centered Architecture Design.

Research is needed to develop architectures that will facilitate the solution of important problems which are not being solved today because our fastest machines are incapable of solving them. This does not simply mean that all that is needed is faster components. On the contrary, the benefit of a supercomputer to users is bounded by the weakest part of the total system in which it is located. The entire design of the computer: the interaction of the input/output, memory and processing systems, the level and types of instructions used and the suitability of the instruction set to the problems to be run on the computer are all important considerations in determining the speed of the system.

Architectural development of very high speed computers therefore needs to be driven from the requirements of the problems which need to be solved. This can only happen in an atmosphere where physical scientists and computer scientists are working side-by-side to understand the joint interaction of algorithm development and computer design.

3. The Need to Build Prototypes.

The least expensive of the three components of a successful research program in high speed computers is the "paper design". The most expensive is the construction of a physical prototype which is vitally necessary to the research program.

"Paper designs" can specify many different architectural approaches, some of which may be impractical to build, hard to manufacture, or impossible to use. Only in actually building a prototype is the system design proven. Additionally, in building and using a prototype, research must progress in several areas. These areas are: the algorithms that specify the problem solutions; the languages that are used to state the algorithms; the architecture that processes the instructions; the circuits and devices that ultimately implement the computations. We feel that success in each of these areas is vital to achieve high performance for the computing machine as a whole. We see a need to consider the computer as an integrated system, and the corresponding requirement to pull together many disparate talents. We feel that the construction of prototypes provides the framework under which such interaction can flourish.

4. Experimental Facility.

One reason for the decline in high speed computing research is the lack of experimental laboratory systems. A laboratory system is one where the machine (or group of machines) and associated software are regarded as a research environment. For instance, the operating system software of a laboratory computer system can be modified to enable the path of interesting phenomena to be traced. The compilers in the system can be instrumented to determine when certain constructs occur; the computer system can be dynamically reconfigured in response to research questions. A laboratory system is therefore one in which experimental computing research is the object of the activity.

An experimental, laboratory system for high speed computing research is a Cray or Cyber-205 class machine. Clearly placing such a machine in an experimental, nonproduction facility is expensive. But without such a machine on which to test algorithms against which to measure the prototype system as it develops, the high-speed computing research has no comparison point, and the researchers have no experience with state-of-the-art high-speed computing.

Such a machine might be placed in a part production, part experimental environment. In the early days of computing, almost all university and laboratory computer centers were organized as combinations of laboratory and production systems. Indeed, John Von Neumann, was a principal user of the machine he created. Today, however, centers which jointly support both users and researchers are usually characterized by a great deal of tension. Since the users are usually paying for the center, the researchers are often eventually squeezed out and subsequently confined to using their own inexpensive minicomputers. Many examples of such situations have led to the current state of high speed computing research where no major university is either developing or has access to a Class VI computer for experimental purposes. As a result, there is a vacuum.

Today, experimental computer science with high speed computing systems is almost non-existent at universities in the United States. Over the past several years, there have been few innovations in high performance computers other than those provided by the development of improved integrated circuits.

If such experimental facilities are included as part of high-speed research programs, we envision a renaissance in high performance computer development with the re-establishment of close working relations between sophisticated users of massive computing resources and the experimental computer scientists who are attempting to discover innovative approaches to radically improved, high performance systems.

5. Summary.

Our country needs very high speed computer research for several reasons. Some of these reasons have nothing to do with scientific research; they deal with issues of prestige and security. Two reasons, however, are entirely scientific. We do not yet understand the limits of such machines. We expect such research will uncover fundamental principles that will allow such machines to be built. To the extent that such research investigations are successful, the ability of other scientists (physicists, chemists, and geologists, for example) is enhanced to solve problems that can not be solved with the computing machines now available. And these problems, the conservation of energy and the productivity of manufacturing, for example, penetrate every corner of our technological economy.

The three-component program (design, prototype, and experimental facility) specified above is not low in cost. However, now is the time to invest in this critically needed research. A viable research program should establish several high speed computing research centers. Each center should have strong academic as well as national laboratory ties. The centers should be fully funded for the three-component programs so research in this vital area can progress quickly.

3. What are the current status and future prospects if federally and privately funded computer R and D?

4. Is the private sector's research directed toward development of computers that will be needed for forefront research being pursued in DOE programs?

I shall respond to these questions together. There is considerable activity in computer R and D in both federally supported work and in the private sector. Because of defense needs of the federal government and business requirements of the private sector, the strongest emphasis has been on information handling, communication, and automation. There has been a long hiatus in the advancement of computational power for numerical problems; the efforts of Seymour Cray providing a notable exception. Now there appears to be a growing wave of interest in computational tasks, driven in part by computer aided design (CAD) systems. and the private sector shows more enthusiasm. There is urgent need for a corresponding renewal of federally supported R and D in the technologies for designing, building, and utilizing new generations of supercomputers. The Department of Energy has the research missions which will need new generations of supercomputers and the research staff to pursue these developments.

Mrs. LLOYD. Thank you Dr. Lester.

Mr. Haussmann, we look forward to hearing from you now.

[The biographical sketch of Mr. Haussmann follows:]

HAUSSMANN, ALFRED CARL, JR.

Associate Director-at-Large, Lawrence Livermore Laboratory, Livermore, California

---

BIRTH Geneva, Ontario County, New York, August 26, 1924

FAMILY Wife: married in 1947 to Alice Wilkerson, Colonial Heights, Va.
Children: Barbara Haussmann Wilson and Mary Haussmann Boeck

ADDRESS Residence: 4261 Drake Way, Livermore, California

EDUCATION Began Hobart College, Geneva, New York, in 1942; selected to attend United States Military Academy, West Point, N.Y.; graduated in 1946 with B.S. in military art and engineering; remained in Army and studied physics at California Institute of Technology, Pasadena, Calif., the U.S. Naval Postgraduate School, Annapolis, Md., and the Pennsylvania State University, University Park, Penn., earning a Master's Degree and doing additional graduate work.

POSITIONS Nuclear Supervisor at Armed Forces Special Weapons, Sandia Base, New Mexico, 1951-1952

Physicist team member on the Matterhorn Project with Professor John Wheeler at Princeton University, 1952-1953

University of California, Lawrence Livermore Laboratory, Livermore, California:

1953-1955	Army Officer invited to work at LLL on weapons designs
1955-1959	Physics Group Leader (civilian)
1959-1962	Division Leader
1962-1966	Associate Director for Military Applications
1968-1971	Associate Director for Plans
1971-1975	Associate Director for Plans and Lasers
1975-	Associate Director-at-Large

MEMBERSHIPS AND ASSIGNMENTS

Member of Navy Special Task Group (Polaris System Advisory Group), 1958-1961
Consultant to Air Force Scientific Advisory Board, 1958-1961
Member of Air Force Penetration Program Panel, 1961-1962
Member of Department of Defense Strategic Systems Committee, 1961-1965
Atomic Energy Commission Senior Reviewer, 1961-1967
Consultant to Air Force Space Systems Division Advisory Group, 1962-1965
Consultant to Air Force Ballistic Systems Division Advisory Group, 1966-1967 (and miscellaneous prior times not documented)
Member of Strategic Intelligence Panel, 1965-1973
Consultant to Air Force Space and Missile Systems Organization Advisory Group, 1967-1975
Member of Defense Science Board Tactical Warning Panel, 1968-1970
Chairman, Defense Science Board Tactical Warning Panel, 1971-1972
Consultant, Defense Science Board
Member of AEC Laser-Fusion Coordinating Committee, 1972-1975
Member of DOD High Energy Laser Review Group, 1972-1975
Member of Union Carbide Enrichment Technology Advisory Committee, 1983-

MY ADDENDA OR EMPHASIS

Education: B.S. - U. S. Military Academy - 1946
(All-American Lacrosse player)

Attended the Naval Postgraduate School
Annapolis, MD - 1949-50

M.S. - Pennsylvania State University - 1952

Additional Highlights: An Associate Director of Lawrence Livermore Laboratory since 1962.

As Associate Director for Military Applications (1962-1968) was responsible for the development of the nuclear warheads for the entire Polaris/Poseidon SLBM Force and the entire Minuteman ICBM Force

As Associate Director for Lasers (1971-1976) was responsible for the growth of the LLL Laser Program from a relatively small $6M per year research program to a $25M per year major National program.

Dr. HAUSSMANN. Although we didn't coordinate it beforehand, I thought Dr. Lester was right on target.

I thought I would speak informally from a few VuGraphs and as usual this program is out of focus.

When I was running a thermal nuclear program a couple of decades ago, I used the following cartoon which sort of combined fact and fiction.

Of course, as you know, the nuclear weapons program of the country through some amount out of good planning and some amount of fortunate circumstances which I outlined in my prepared statement, got into using modern computers at a very early date.

The second bullet there sort of illustrates part of the problem with getting on with new supercomputer technology. Since it is the leading edge, any given model will run between 10 and 100 units and it is not exactly the place the big companies put there both their monetary and personnel resources. So there aren't too many players in the game.

Just as a matter of bemusement you can look at the last bullet there. I would hate to do my problems that way.

I won't dwell on this too much. It has been mentioned so many times but more and more the leading lights in the universities and national laboratories are realizing the magnitude of the bridge that supercomputers are between theory, that is understanding of the science and technology involved, and actual real life experiments. Then, of course, once there is a good correlation there you can talk about extrapolations forward—things you haven't done experimentally.

You can just use much more theory and put in much more geometric detail if you put all this theory into the supercomputers. That is something the weapons programs have known for decades and belatedly is becoming more widely known in the universities.

I think if they had had the appreciable quantities of leading edge computer time that they should have had this would have been a discovery made decades ago.

This is an answer to your first question. It is really a hedge rather than an answer. The reality of the situation is contained in the first bullet and that varies from program to program. I list some examples. From Livermore, for example, the weapons program spends about 10 percent, 9 percent, of its dollars on essentially buying computer time and a comparable amount of dollars buying people to run problems on the computer. In other words, about 20 percent of its effort.

The laser program, it lumps laser fusion and, that is, inertial confinement fusion and laser separation together, but probably the inertial confinement fusion uses about 5 percent of its dollars and, again, a comparable amount of effort on people running the problems, for a total of about 10 percent.

We run, as you probably know, the National Magnetic Fusion Center, and that represents about three percent of the magnetic fusion budget.

I would hazard the guess that you will see that go up several fold in the next few years and the reason is that large machines are getting much more complex and there is still plenty of room for the magnetic fusion people to model in the same amount of detail, for example, as the weapons programs do. So that would be a prediction on my part.

What one can say about all these programs, in my opinion, is that as much computer time as they utilize is less than they should be utilizing. Then that should be compared with a large fraction of the community that doesn't utilize computers at all. That is a combination of a historical situation—that is the way they developed management attitudes and budgetary constraints.

I think, again, when the national universities are turning out people familiar with large-scale numerical modeling on modern computers you will see an absolutely dramatic change in the utilization of large-scale computers and a wide range of commercial and national problems.

DOE has been very successful, particularly in the fifties and sixties, in providing market pull for commercial computer outfits to provide leading edge computer products and in using the computers. The market pull went away in the seventies because the DOE's and its predecessors were not putting the moneys in that they had earlier and, relatively speaking, the amount of moneys they were putting in became a very small fraction of the market, after all, the information handling market today is probably well over a hundred billion dollars a year.

When people talk about market pull being the prime solution to supercomputer progress, I question whether the Government will put in the resources to really accomplish that.

I will come back to things we should do in a while.

There it is actually on the bottom of that Vugraph. I think all of these organizations—industry, academia, the National Labs, the commercial sector that might actually make these machines and then use them, all should, particularly where they have a position of strength, they should be allowed to show what they can do.

I have read the prepared statements of a number of people and listened to some of their remarks and one point I would really make is that all of these organizations have to work together. You can't just focus on just one or two items and expect to run a good race, for example, with Japan, Inc., which does use all of these techniques.

As you are probably aware, at Livermore we have probably the largest national lab supercomputer design project and the thing that we would propose to do is show some of these leading techniques and work hand-in-glove with universities and industry so they could make good use of the best of what comes out of our operation, and we could learn from them.

It has been mentioned before the need for present generation commercial supercomputers and that is certainly true. There is some new ones coming along. The computer centers that exist around the country should be upgraded, probably some new ones started, and also the capabilities expanded.

For example, a possible near mark term solution for academia, and energy research, would be to expand the Magnetic Fusion Computer Center by a factor of two or more and provide on a rather short time scale computer time for a range of organizations, universities and energy research.

In the longer term you could look into other techniques to handle that but that is probably your fastest response capability.

Mr. HAUSSMANN. I guess I should note right here today that I was pleasantly surprised at the prepared Cray research statement expressing optimism in getting one hundredfold or two hundredfold improvement over the Cray-I by the end of the decade. I wish them luck and I hope they succeed. Obviously, this VuGraph was considerably more pessimistic with regard to the private sector as it is set up now.

There are very few players and they have a real millstone around their neck in the sense that already the Japanese—the basic technology that supports making large-scale computers, that is, the component technology—the Japanese are already out in front of the U.S. base technology.

There seems to be some question about that but I firmly believe that. Not only that, you can't buy the leading edge Japanese technology in America because they retain it for utilization of their own industries.

Even so, the items you can buy are quite competitive and sometimes better than the U.S. counterparts.

Again, projects like the Livermore project should be melded into a national program wherever there is a strength to build on, whether it is national lab ability to turn out software algorithms or experiment with leading edge hardward, that should be melded into the national program.

And communications or technology transfer is a very important ingredient of that. Again, with the universities, private sector, national laboratories .

Mr. HAUSSMANN. We want to get what we can out of the private sector and encourage them to accelerate their pace of improvements on computers. That is crystal-clear. They certainly will be the organizations that provide the hardware for expanded computer time for the organizations that you will want to provide that time for.

You simply can't rely on market pull, in my opinion. I notice other people that will testify today disagree on this but, again, I think you have to use all segments of strength in the computer society to maintain the national leadership.

Illustrative of the problem you have is here is a proposal by a Japanese supercomputer outfit for their SX-1 and SX-2 computer which will be available in Japan early next year. They haven't yet announced when it will be available in the United States or elsewhere.

If that computer meets the characteristics they are talking about, it will be better than any computer that is expected to be out in a comparable timeframe from U.S. manufacturers. That would be a major step forward for the Japanese and a target for us to take on.

Unless we take some of these steps that have been talked about by the various speakers, the second half of the eighties is going to be a very traumatic experience for the U.S. leadership role.

A main difference in why this is coming about, incidentally, is Japan, Inc. does one thing that just straight U.S. market pull will not do. They tend to plan for the longer and strategic haul rather than the shorter haul or tactical haul that may enhance near-term profits. That is the thing that we have got to counter.

Lowell, do you have that last cartoon?

I hope everybody gets properly motivated. You can be too complacent on occasion and it has happened in the past.

I would be glad to answer any questions that you have.

[The prepared statement of Dr. Haussmann follows:]

VERY HIGH PERFORMANCE COMPUTING:

National Needs And Options In Energy R&D

Prepared Statement And Testimony Graphics Of
A. Carl Haussmann,
Associate Director-At-Large
Lawrence Livermore National Laboratory

Before The
Committee on Science and Technology
U.S. House of Representatives

2318 Rayburn House Office Building
2:00 PM, 15 June 1983

In order to put utilization of very high performance computers into proper context, I will speak a bit more broadly than to just the immediate energy research focus of this hearing. In particular, since I have been heavily involved with computing during practically its entire modern "electronic" era, I will briefly recapitulate its history from *my* perspective. It may help illuminate the needs and directions optimal for future progress.

Modern supercomputing, at least as I know it, began a bit more than a third of a century ago under an interesting, and indeed quite remarkable confluence of circumstances. John von Neumann, a world-famous and incredibly gifted mathematician, became quite interested in what for that era was very large scale supercomputing. (By today's standards, the computing hardware which he developed and used ranks below that of the most modest microcomputers - on-a-chip.)

Von Neumann not only developed a computer which was quite advanced for its time, but very importantly, began to apply this new capability to what for that era constituted large scale modeling of the physical world. In the same time frame - the early 1950s - Eckert and Mauchly were designing, and their computer company was fabricating in Philadelphia, the first commercial electronic computer, the Univac I. In this same period, the AEC was inaugurating its thermonuclear explosive program. It is not generally known, but the detailed thermonuclear component design and yield calculations for the first major American hydrogen bomb were accomplished at Princeton's Project Matterhorn under the leadership of the noted physicist, John Wheeler, and his brilliant young protege, John Toll. Both of these men were of the same mold as von Neumann and likewise realized that the ability to apply the maximum amount of theoretical understanding in the greatest geometric detail to these extremely complex design/yield calculations could only be done by using the most modern electronic computers. Hand calculations using mechanical desk calculators were totally inadequate. The upshot was that major portions of the early generation thermonuclear weapon designs and test series (1952 and 1954) designs were accomplished right on the floor of what had by then become Remington Rand's Philadelphia computer production plant.

It was from this interesting confluence of critical circumstances - key individuals, challenging programs, and evolving computer technology - that the AEC's capability and interest in modern computing derived. They rapidly spread to the weapons laboratories, Los Alamos and Livermore, both because of the obvious opportunities presented and because, when early in 1953 the thermonuclear explosive design portion of Project Matterhorn was dissolved, some of its personnel migrated to those organizations. From this "seeding" these two Laboratories have remained preeminent in numerical modeling, always exploiting the best computers as successive generations become available.

The laboratories' interest, the AEC's funding, and the private sector's dedication led to advances in scientific computer capability of a factor of 100 in the decade of the 50s and similar advances in the decade of the 60s. Since the AEC's interest and funding were the primary "market pull" of those times, the most capable computer design talent of the commercial sector was dedicated to and was quite adequate for the creation of one generation after another of "leading edge" scientific supercomputers.

By the early 1970s, however, the situation had changed dramatically. The AEC and its successors (first ERDA, and then DOE) were no longer providing the dominant "market pull". The world-wide commercial sector was expanding its use of small computers at an enormous pace and the weapons laboratories were unable to obtain supercomputer funding levels which could support a vigorous acquisition program of "leading edge" computers. Virtually all of the talent within computer design and manufacturing organizations migrated to where the money was, the mass utilization market. As a result, the growth in supercomputer capability during the 70s was very modest indeed, a few-fold.

The reason for slowed supercomputer capability growth since the early 70s was not predominately a deficient technology base. Others, notably the Japanese, have observed this situation. They are now targeting several prime areas to gain a preeminent position in this field of extraordinarily high leverage. They recognize that supercomputers may well be a key to dominance in large scale science and technology and in medium to high technology industries.

First, they have moved to advance the supporting technology base for supercomputing distinctly beyond that of their American counterparts. The impact of this can be observed in their aggressive and successful take-over of the market for high density memory chips, perhaps the single most critical element of modern supercomputers. Their technology base efforts of the late 1970s also gives them a factor of several-fold greater inherent speed in their fastest integrated circuits. Their commercial-grade technology is now more advanced than that of the U.S. The U.S. mass computer market does not require fast logic components and high density memory chips as do state-of-art supercomputers. It is also Japanese practice to retain their best components for their own supercomputer manufacturing operations.

Secondly, the Japanese have noted that, since U.S. computer development is dominated by "market pull," U.S. computer corporation planning tends to be for the "next few years" and not for the long term. On the other hand, "Japan Inc.," the well-known consortium of Japanese government, industry, and academia, has targeted the supercomputer arena for the longer haul. This targeting is especially effective because of the relative "vacuum" of ambitious U.S. supercomputer efforts and corresponding technology base development. Japan is further aided in its national priority pursuit of large scale computers by a clever marketing plan. It is turning the decades-long, ongoing, full-service efforts of major U.S. companies such as International Business Machines and Digital Equipment Company to their advantage. By making their machines compatible with U.S.-developed software, Japanese computer companies do not have to provide the mass of software support for which U.S. organizations continue to expend great effort and resources.

Another relatively recent Japanese targeting action is to move toward "breakthrough" advances in supercomputer capability, 1,000-fold capability gain by the end of this decade. They are also going for revolutionary advances in techniques to utilize this huge new computer power effectively. In my opinion, their goal is to dominate the world-wide information market. This imaginative and thoughtfully considered "game plan" is likely to win the preeminent position in the emerging "information society" for them unless

U.S. government, industry and universities shuck their "middle age spread" and posture themselves to "run the race" that the U.S. is quite capable of winning when highly motivated. A powerful and coordinated U.S. effort is necessary; a "business as usual" approach is doomed to failure!

To date, the required U.S. motivation is only partially discernable. Academics note that even today availability and utilization of leading-edge computer capability on campuses is almost nonexistent; with respect to supercomputing, the universities, generally speaking, are in the "Stone Age." This is a National disgrace. If the universities were turning out graduates with appropriate levels of large scale numerical analysis expertise, a much greater utilization of large scale numerical modeling by U.S. industry would be seen. There are many situations where such modeling is outstandingly cost-effective. This will become even more so as computers become much more capable and much less expensive, and as general technology utilization becomes more widespread. (The Japanese make no secret of their realization of these basic facts.) Academics also correctly observe that large scale numerical analysis is an absolutely necessary bridge between science and engineering theory (for which I would use the term "understanding") and real world experiments of all kinds. Today's situation is very similar to what it was for von Neumann and Wheeler over three decades ago. The use of supercomputers to numerically model the physical world provides the maximum opportunity to utilize theoretical understanding to the fullest complexity and value when making comparison with the real, experimentally observed world.

For the United States to move out in the expeditious fashion necessary to maintain National computing preeminence and to generate the required rate of improving computer capability, several things must be done. A coordinated, long-term plan and National level oversight, preferably with a centralized management structure, need to be set up. Key government agencies (DOE, DoD, NASA, NSF) must be well-coordinated and must also "pull their own weight" with respect to effort and resources. Where the National laboratory system has unique, relevant capabilities, they should be vigorously utilized. Arrangements should be made which would allow the private sector to make its maximum contribution. The academic community must be facilitated to join today's and tommorrow's exploitation of the supercomputer world. Realistic goals and priorities must be clearly stated.

Existing plans are not, in my opinion, realistic. They seem to stress modestly increased inter-group communication and limited additional funds, relying primarily on the "market pull" mechanism. This level of response is inadequate to maintain our National computer preeminence. "Market pull" alone is very unlikely to be a satisfactory solution at any resource level the Nation would be willing to commit to. The increased resources the DoD (DARPA) is planning to spend in congruence with DARPA's specific institutional interests are targeted for such things as gallium arsenide technology base improvement and artificial intelligence developments. While these technology thrusts, if successful, can be of major long term National value, they are unlikely to be of major benefit for the short and intermediate terms.

The DOE's present "market pull" proposals appear to be completely insufficient to regain the supercomputer initiative. DOE's intentions with respect to significantly increased resources in computing hardware and software R&D within its laboratories and the universities are not readily discernible at present.

Congress has, over the last few years, made some preliminary initiatives in support of supercomputing, but to date they have not been "full goal" efforts, nor have they been vigorously funded. They certainly have not simulated effective responses from the bureaucracy.

Organizations such as the Livermore Lab have begun to provide specific, action-oriented inputs to those individuals and officials interested in the subject of supercomputing, supercomputing utilization, and National technological leadership retention. Our current package along these lines is a paper, LLL-PROP 185, with attachments, entitled "Supercomputing Research and Development: The LLNL Program and Proposal," dated March 1983. Many interested parties already have copies of this package. It is available to others by contacting Betty Dolen of my office, Lawrence Livermore National Laboratory, P. O. Box 808, Livermore, CA 94550, (415) 422-4855. This package has been provided to DOE (Office of Energy Research Director, Al Trivelpiece and Assistant Secretary for Defense Programs, Herman Roser) and to DoD (DARPA Director, Bob Cooper). While this package does not constitute a National supercomputing "game plan", it does contain major components potentially useful in formulating such a plan.

An important ingredient in binding organizations together in this goal of maintaining supercomputer preeminence would be multi-directional communications brought about by "technology transfer" between contributing National laboratories, leading edge commercial ventures, and academic contributions. We at Livermore are stepping up our already significant contacts with other organizations in these areas and intend to accent constructive communications with them.

Fusion energy is an example of DOE energy research which is likely to benefit substantially from greatly increased (factors of 100 to 1,000) computer utilization. (Realization of such capabilities is a technically and economically feasible goal during the coming decade, as the Japanese have noted.) We at Livermore believe we can attain such milestones with the our existing supercomputing team and goal structures. To support such efforts, implementation resource levels will have to be increased by 10 to 20 million dollars a year. As a preamble to one of our Lab's thirtieth anniversary monthly publications (Energy and Technology Review of October 1982), I addressed just this issue. My comments are appended to this prepared statement.

I strongly endorse Congressional encouragement to DOE's Energy Research Office for immediately providing adequate quantities of "leading edge" computer capability to energy research activities nationwide. Expansions of facilities such as the National Magnetic Fusion Energy Computer Center network are reasonable ways to attain interim capability reliably and at an early date. Such approaches will introduce to a distinctly wider set of individuals and organizations the benefits large scale computing can bring to energy research, both to longer range (MFE and ICF) and intermediate range projects.

A supercomputer application that could be economically crucial, but generally has had little attention focused on it, is the commercial utilization of supercomputers to design and build the following generation supercomputers. Livermore's S-1 supercomputer design and exploitation project is perhaps the major U.S. example of what the future holds for high technology, fast-moving, commercial efforts. Through unprecedentedy comprehensive creation and utilization of computer-aided design and manufacturing techniques, and exploitation of very high performance and high density memory technologies, our small team, made up of premier quality people, has very rapidly implemented supercomputer design concepts into working hardware in two generations of S-1 supercomputers. The fundamental lesson of the S-1 effort for American computer manufacturers, however, is that it is feasible to assemble a small team of very high quality. Thus total manpower and overhead costs and delays in creating supercomputers can be slashed drastically and the design and manufacturing cycle times, along with the relevant costs, for state-of-the-art computer systems can be reduced several-fold, relative to present U.S. commercial experience.

I mention our S-1 project for two reasons. The first is that the existing S-1 capability almost certainly represents the Nation's most sophisticated and capable supercomputer design and prototyping effort and the one with the greatest ability for rapid response. It is ready, today, to go "head-to-head" with any international competition. It should therefore have a major role in any accelerated National program. More importantly, the general concept of massively computer-supported design and manufacturing will surely come to be utilized by most (successful) medium and high technology, rapidly-moving industries during the rest of this century. There is no reason these same basic techniques that the S-1 group applies to computers can't be applied to designing airplanes, skyscrapers, or refineries. Believe me, those organizations, American or foreign, which most effectively develop and exploit these techniques will dominate their fields of endeavor.

Preamble - Energy and Technology Review - October 1982

This issue of Energy and Technology Review is the second of two commemorating the 30th anniversary of the Lawrence Livermore National Laboratory. The first issue reviewed a number of the Laboratory's original programs. Here, we review our major energy programs, one of which (magnetic fusion) has been a Laboratory responsibility during the course of our entire existence. Several of these energy efforts hold the prospect of great value to the nation and the world. Interestingly, some of them (e.g., the fusion programs) are of such complexity and duration that they will not yet be fully integrated into the country's economy at the time of our 60th anniversary and, hence, may still be topics of discussion during that celebration.

The first articles describe the Laboratory's fusion programs, magnetic fusion and inertial confinement fusion, which are alternative and independent routes to the almost limitless energy source that fusion promises to be. Magnetic fusion energy, in which a dilute fusion fuel is confined with magnetic fields, is the older of the two programs, and it has made major strides toward understanding the science involved. This country is dedicated to the pursuit of magnetic fusion energy; its current level of support is nearing $500 million annually.

Inertial confinement fusion, in which pellets of fusion fuel are compressed with intense pulses of energy, is a more recent program. It too has made considerable progress in the 20 years of its existence and, of course, is known to work at large (nuclear explosives) scale. The inertial fusion program has dual goals: for the near-term, to elucidate some of the physics associated with nuclear explosives, and, for the long-term, to become a candidate for commercial fusion energy. Current national support, which is approaching $200 million annually almost totally is dedicated to the basic understanding and demonstration of inertial confinement fusion at laboratory scale and to military research. The major emphasis in the article on inertial fusion in this issue is, however, on the advances in basic understanding and on commercial power prospects.

Both fusion programs are accumulating the understanding required for fully credible scientific feasibility demonstrations. Although neither demonstration has yet been made, both approaches have excellent prospects for accomplishing this in the next decade or two. It is important to remember that even after scientific feasibility is demonstrated, the resources required to build a reliable and economically competitive power plant are undoubtedly many times greater than the resources required for scientific demonstration.

There are some real challenges to such large-scale, long-term efforts as magnetic and inertial confinement fusion, especially in this era of concern over balancing government spending and income. This budgetary issue suggests that near-term research and development in major programs probably will be done in an environment of relatively level resources. Progress under these circumstances is especially difficult, as the major experimental facilities tend to cost from a quarter of a billion to more than a billion dollars. Additionally, these facilities take nearly a decade to execute: three years or more to introduce as part of the national plan and five years to construct. Obviously, it is important both to minimize the number of such major facilities and to pursue the programs and their facilities in a fashion that maintains national interest at all levels.

That is the challenge presented to the managers of these high-leverage, large-scale, long-term programs. The communities involved must, and I strongly believe can, rise to the occasion. As programs strengthen their scientific bases of understanding, they become more insightful and sophisticated in the further pursuit of such knowledge. As programs face challenges such as those noted above, they provide a remarkable level of stimulation to the ingenuity and inventiveness of the people involved.

Today, there is also the prospect of really major expansion in the utilization of numerical analysis. Very probably for at least the next two decades, the large-scale electronic computers that support detailed numerical modeling of major scientific endeavors will increase in capability by a factor of a hundred each decade. This increase should provide major programs with the opportunity to do far more detailed anaylses than have ever before been possible. It will be interesting to follow these programs as they exploit this pending explosion in computer capability for the furtherance of their goals.

The Laboratory's energy and resource programs are reviewed in the third aritcle in this issue. While the potential value to the nation of these programs is also great, they have quite different characteristics from those of the fusion programs. The science and technology of these energy and resource programs can be demonstrated in much shorter times (perhaps a decade) and with much smaller financial resources (perhaps tens of millions of dollars annually). Of course, integration of successful programs into the country's economy will take considerably longer because of capital funding limitations and historical constraints.

A prime example of a program of great potential that can be investigated at a "bargain" price is the Laboratory's in-situ coal gasification research. If successful, it would increase the already immense economically retrievable coal reserves of the U.S. several fold. While the need for additional gas supplies for the next decade or two may not be large, this program could supply enough gas that when converted to liquid fuels (e.g. methanol) could provide independence from foreign oil supplies at economically attractive prices. It is difficult to understand why this nation is not expeditiously pursuing such a high-potential and low-developmental-cost program.

Within the context of its charter and capabilities, the Laboratory is studying possible new efforts that would be of great value to the country and the world. Some of the new efforts undoubtedly will be reported on the occasions of future anniversaries. When implemented, these are the programs that will make it feasible to increase the productivity and standard of living of the U.S. and the world within a secure environment.

Carl Haussmann
Associate Director at Large

What Are Supercomputers?

- Top-of-the-line digital electronic computing systems
- Typically only several dozen in existence at any given time
- 'Cottage industry' products in the U.S.
    - Relatively small economic potential
    - Few people-and-companies know how to create them
    - Few people-and-organizations have had experience using them
- Unit costs are 10 - 20 M (1983) $
- Present ones each compute about as fast as entire human race could, if completely equipped with top-of-the-line pocket calculators

Why Are Supercomputers Important?

- Physical modelling
  - Detailed view of the physical universe
  - Sensitivity to assumptions easily studied
  - A "bridge" between theory and experiment
- Perform 'experiments' inside computers
  - Solve laws of nature subject to initial assumptions
  - Generate 'histories' of snippets of model universes
  - Far cheaper and quicker than doing real experiments
    - Best ones are comparably informative
      - Assumptions usually imperfect
      - Powers-of-observations usually better
- U.S. nuclear weapons program critically dependent on supercomputers since its inception
  - Infeasible to execute sufficient experiments
  - Diagnostics of experiments are insufficient
- Magnetic and inertial confinement fusion programs have largely adopted weapons program computing techniques
- Other energy R&D programs would benefit by doing likewise
- Constitute the 'leading edge' of all computer development
  - Enabling technology for all lower-performance systems

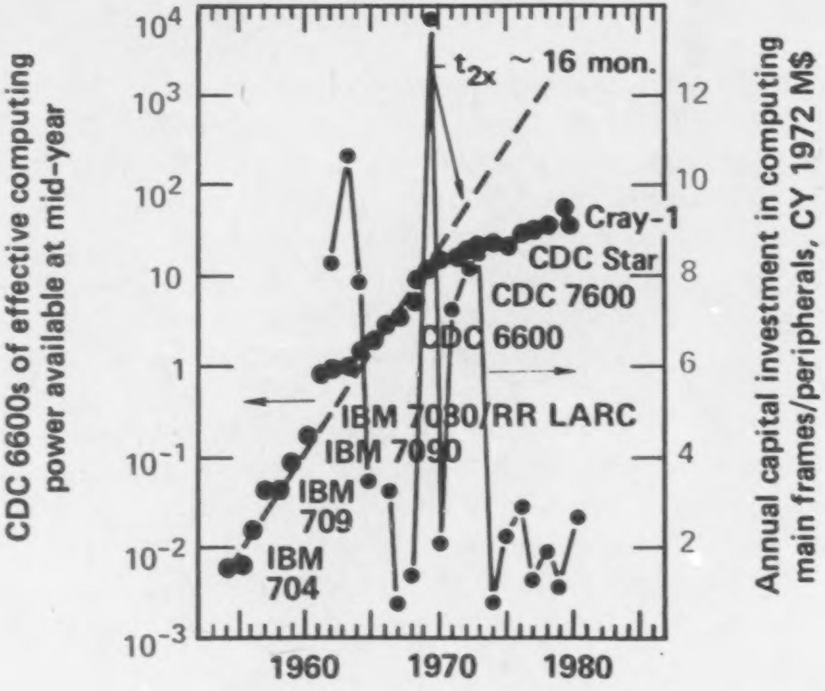
What's A Reasonable Supercomputer Usage Level?

- To whatever point the marginal dollar is most effectively spent, e.g.,
    - On "experiments" performed in computers vs. on real-world experiments
    - On computer-centered experimental data storage-analysis-and-control vs. on alternative means
- LLNL weapons program spends about $22 M of its $235 M budget on the central computing facility, or about 9%
- LLNL laser programs (ICF and LIS) spend about $3.5 M of its annual $110 M budget on the central computing facility, or about 3%
- National magnetic fusion energy program spends about $15 M of its $466 M annual budget on NMFECC, or about 3%
- Cost-to-programs of the computational efforts of professional researchers is comparable to the computer center costs
- Energy Research programs could productively triple their computer usage, following the highly successful example of the DoE weapons program

DoE As A Buyer/Developer of Supercomputers

- DoE community most experienced in all aspects of supercomputers
  - Four decades of weapons program development/exploitation of the entire set of underlying technologies
    - Hardware
    - Systems software
    - Scientific algorithms
    - Applications programming
  - More recent experience by inertial and magnetic confinement fusion communities, high energy physics community, etc.
- DoE highly successful as supercomputer buyer when it had the only resources, during the 1950s and 1960s
- Very unsuccessful 'market-puller' during the past dozen years
  - Available resources dwarfed by those of business sector
  - No reason to expect changed situation in the 1980s
- DoE should:
  - Support vigorous "technology transfer" between industry, academia, DoD and its own areas in scientific supercomputing
  - Buy present-generation supercomputers for its currently computing-impoverished communities
  - Develop next-generation supercomputers for the impending needs of its computing-intensive communities

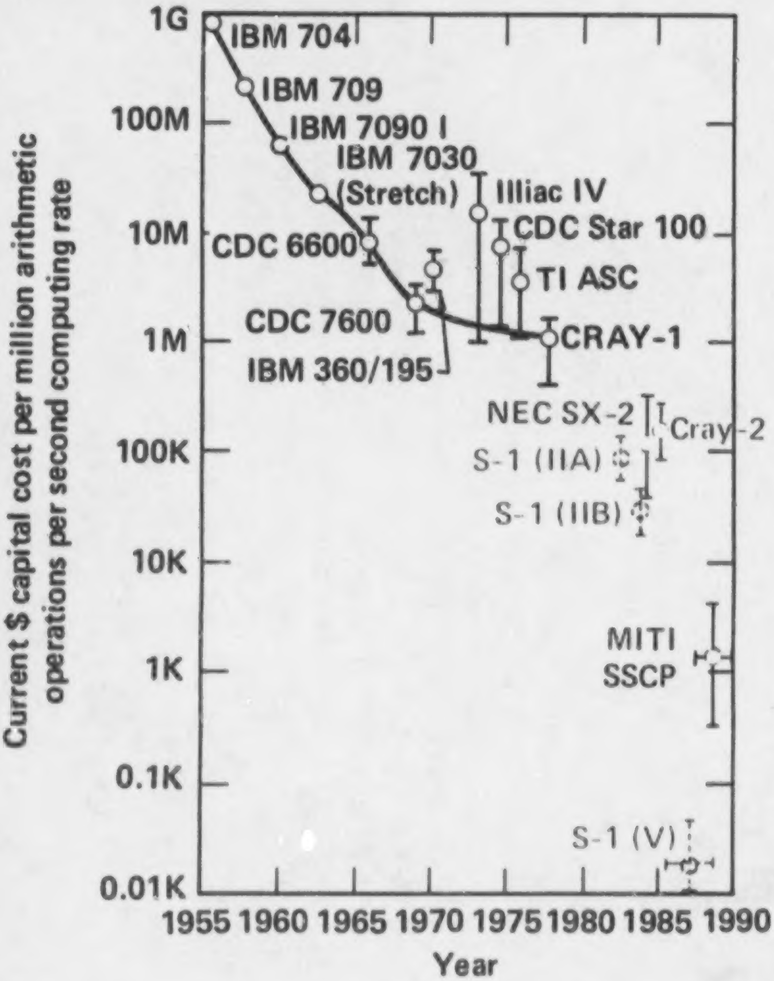
CAPITAL COST AND PERFORMANCE HISTORIES OF THE LLNL COMPUTING FACILITY



Federal And Private Supercomputing R&D

- U.S. private sector status and prospects
  - Two major companies
    - Staying power of one is questionable
    - U.S. tech base erosion plagues other
    - Course and pace of future development uncertain
  - All other U.S. private sector efforts much further back
    - Durability of recent interest is unclear
- Federal efforts status and prospects
  - Accelerated DARPA effort focussed on artificial intelligence
    - Potentially very important
    - Impact on existing DoE and DoD computing needs is negligible--symbolic vs. arithmetic computing
  - LLNL's S-1 Project (co-sponsored by Navy and DoE) is only Federal scientific/arithmetic supercomputing endeavor
    - Emphasizes DoD/DoE-oriented innovation, rapid turn-around and low cost
  - Best U.S. near-term bet is to capitalize on S-1 strengths
    - Accelerate hardware development and dissemination of S-1 technologies to industry
- Better coordination of Governmental, private sector and university efforts urgently needed
- Major resource increases required

THE COST-EFFECTIVENESS OF HIGH PERFORMANCE, GENERAL APPLICATION DIGITAL COMPUTERS: HISTORY AND FORECAST



Private Sector Supercomputer Development

- Completely inadequate to meet DoE program needs
    - Last dozen years saw 3-fold capability advance
        - Previous two decades each saw 100-fold increases
    - Modest scale supercomputers presently in commercial development
        - Low likelihood of technical/economic success
- Available technology base 2-4 times poorer than that available to Japanese supercomputer makers
    - Trends indicate that this gap will widen for the next few years
- Both U.S. private sector supercomputer houses have encountered major difficulties in moving to next-generation supercomputers
    - High capital costs slow high tech innovation
    - Erosion of U.S. digital electronics tech base
    - Not changing with the times
- Japanese have announced supercomputers (e.g., NEC SX-2) substantially more powerful than the best under development in the U.S. private sector
    - Japanese long-range planning contrasts sharply with U.S. 'market pull' complacency
    - U.S. assets focussed on low-middle performance 'mass market'
- Outlook is for a U.S. supercomputing debacle in the second half of the '80s

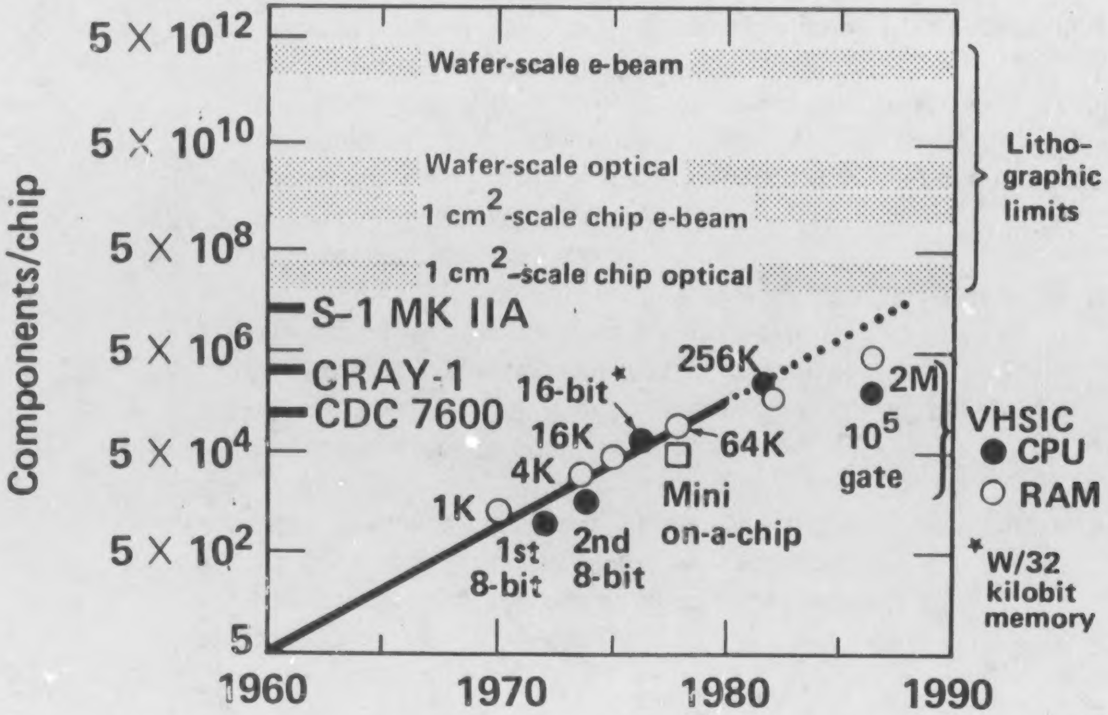
U.S.-Japanese Technology Base Issues

- Asymmetic approaches
  - Technology selection: Japanese emphasize long-term; U.S. short-term
  - Japanese emphasize technology strategy; U.S. optimizes technology tactics
  - Japanese develop tech base cooperatively and exploit it competitively; U.S. emphasizes across-the-board competition
  - Tech base development: Japanese invest thoughtfully; U.S. speculates intensively
  - Japanese seek domination of critical high-tech markets over long haul; U.S. looks for short-term payoffs anywhere/anyway
  - Japanese capital allocation for tech base advances determined by long-term national interests in ever-higher technology; U.S. technologists compete with everyone else for capital
- Tech base asymmetries favor Japanese, in most respects, in the longer term
  - U.S. adaptation of several Japanese approaches is imperative
  - Aggressive all-sector tech base development, improved inter-sector effort coordination, Government-fostered industrial sector cooperation, improved capital/incentives structures

Technology Base Asymmetry Consequences

- Digital semiconductor electronics tech base
  - U.S. substantially broader than Japanese
    - Addresses near-term markets as well
      - E.g., microcomputers
  - Japanese focus on relatively few, high volume components
    - E.g., high-density (64K-256K) dynamic random-access memories (DRAMs)
  - Major U.S. deficiencies at high-performance end
    - 2-4-fold lower speeds and densities in top-of-the-line logic and memory devices
    - Progressive deterioration of U.S. position
- Consequences for U.S. supercomputer capabilities are ominous
  - U.S. supercomputer builders can't buy top-of-the-line Japanese components
  - Second-cut Japanese components generally superior to best-available U.S.
  - Supercomputer 'tricks-of-the-trade' now well-understood
  - Intrinsically inferior U.S. supercomputer posture currently being confirmed, e.g., NEC SX-2
- U.S. inferiority in supercomputing will follow
  - Japanese will dominate a major technological wave of this century

DIGITAL LOGIC CAPABILITY COMMERCIALLY AVAILABLE ON A SINGLE CHIP



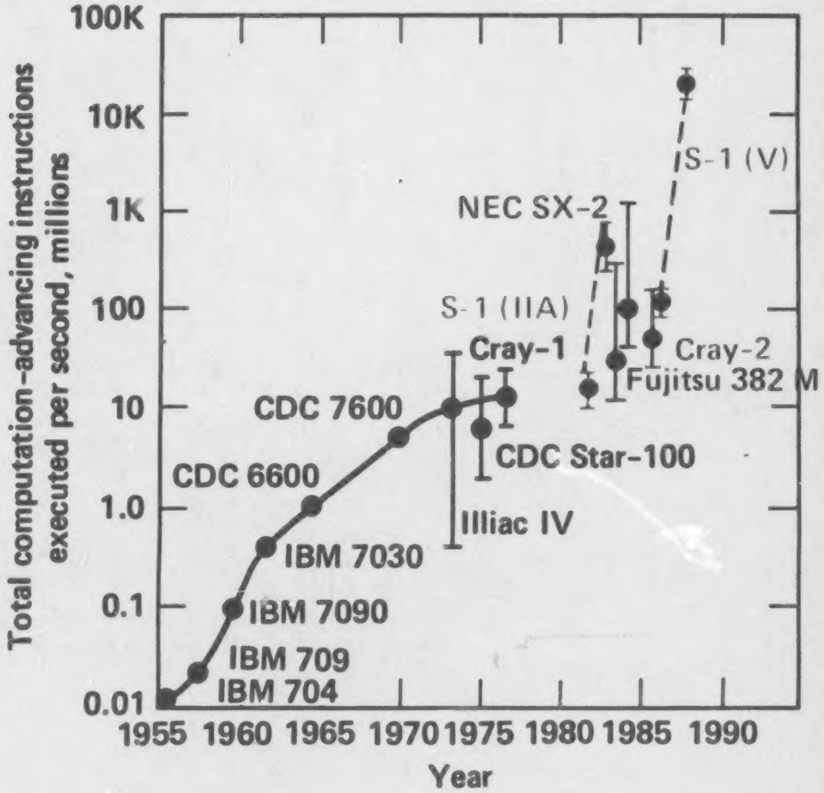
Supercomputing In Academia

- Intrinsically crucial
    - Primary source of conceptual-level innovation in U.S.
    - Essentially unique source of trained professionals
    - Major resource-consuming area in U.S. scientific/technical structure
- Practically non-existent through the present
    - Unit buy-in/upkeep costs are several million dollars per year
        - Large by standards of most universities
    - Long-term commitments necessarily involved
    - Supercomputing research/teaching needs/opportunities foreign to most academic administrators
- Sharp contrast to Japanese situation
    - Government emphasizes immediate availability of best industrial supercomputer offerings to universities
    - Universities are key members of Japan, Inc. team in present supercomputing challenge
- Proper U.S. response is to vigorously foster supercomputing availability to academia
    - Supercomputing resource-sharing from DoE National Labs as immediate corrective step
    - Then steadily build up national research supercomputing networks--natural distribution of capabilities and costs

Attaining 1000X Greater Computing Capability

- Single computers won't get there in 1980s
  - Physical limitations of underlying information-switching devices
- Intelligently federate modest number, N, of fastest-feasible computers
  - Divide computing work among them roughly equally
    - Work division itself is computer-automated
  - Work then gets done up to N times faster
  - Huge (large N) federations precluded by complexity
- Single computers can be expected to increase about 10X in speed by 1990
  - Monolithic implementations may have mass-production unit costs as low as $1 K
  - Federate 64-256 such systems together to get 1000X
  - Federation cost will be less than a single 1983 supercomputer
- Monolithic supercomputer implementation is goal of LLNL S-1 Project in 1986-88 time frame
  - Creating prototypes of single and federated supercomputer systems through the present
    - Current single S-1 computers comparable to fastest in existence
    - Present-generation S-1 computer federation will be most powerful system by far
  - Severely funding-limited rate-of-advance through the present

HISTORY AND FORECAST OF PERFORMANCE LEVELS AVAILABLE FROM LARGEST GENERAL APPLICATION DIGITAL COMPUTER SYSTEMS



Where Does 1000X Fit In Energy R&D?

- Japanese end-1980s goal is 1000X present U.S. supercomputer performance level
- Where might such capability fit in the energy R&D picture?
    - Magnetic fusion energy
        - Advanced plasma confinement/heating systems design-and-evaluation
    - Inertial confinement fusion
        - Quantitative evaluation of competitive driver proposals
    - Fossil fuel combustion
        - Enhanced efficiency, lower pollutant engine designs
    - Nuclear reactor design and safety
        - Faster, cheaper approaches to improved power systems
    - Weapons and weapons effects
        - Enhanced military effects, lower resource costs, less collateral damage
- Required scientific base already in place
    - Faster computing rates and larger memories needed to realize gains
    - Corresponding software and algorithms development/exercise

Salient Features Of The S-1 Project

- Small group of bright, intense young professionals
- Pioneering new approaches to supercomputer design and creation
  - Unprecedentedly high degree of computer automation of work
  - Using hardware and software tools of the late '70s, rather than late '60s
    - Typical of very best of U.S. and Japanese industrial efforts
  - Hardware outputs are 'proof-of-principle' of automation efforts
- Emphasizing rapid turn-around of designs and hardware prototypes
  - Swift capture of new tech base advances
  - Highly adaptable to new exploitation opportunities
  - Inherently low cost
- Focussing on applications needs of Navy and DoE sponsors
  - Unique, 'in-house' degree of responsiveness to test-and-evaluation results
- Emphasis on high-efficiency software engineering practices, e.g.,
  - Transport of high-level language software
  - Use of modern programming languages and environments
- Transferring designs, tools, software and results to American industry
  - Long before Stevenson-Wydler Act
  - Uniquely high leverage--no proprietary limitations
    - Japanese cooperative tech base development flavor

Mrs. LLOYD. Thank you very much.

Dr. Haussmann, I am going to address this first question to you after that cartoon there.

I think you are very familiar with the Lax Report and it stated that lack of in view of maintaining our leadership in the state-of-the-art computers is very slim.

Mrs. LLOYD. Do you agree with this or are you saying that we have already lost it?

Dr. HAUSSMANN. I believe we have lost it. You will be losing it in the next year or so, that does not mean that that has to continue for any length of time.

The United States clearly has the capabilities in the universities, National Labs, industry, to take on any world power and stay out in front. We will have to change our ways to do that.

Mrs. LLOYD. Are the national laboratories working with the computer companies in development of supercomputers?

Dr. HAUSSMANN. At least speaking for ourselves, the following things are what we are doing in concert with the Navy and DOE sponsored S-1 project—at least one of the outputs of that project aside from actual computer designs—is really an advanced software routine that allows a very small number of talented people to iterate a computer design, a major change in computer design, in a 2-year period.

We, for example, that design team is about four people. As a result, that is several fold shorter turnaround time than most commercial outfits have. The Navy and DOE have made those design techniques available to major U.S. computer manufacturers.

The Navy has just sponsored a symposium at Livermore that will actually also evaluate the product turned out by these machines. We have had major computer organizations, for example, the Digital Equipment Corp. and IBM, have been out in mass to see if we can get an industrial participant program going at the laboratories for enhanced technology transfer.

IBM has also offered to make some of their technology available that would be useful to us.

Mrs. LLOYD. Pardon me.

Dr. Roberts and Dr. Lester, how do you feel that limited access to supercomputers hurt or damaged basic energy research? Are there really researchers out there who are really limited because of inadequate access?

Dr. ROBERTS. I think that one can do this by example. I have tried to provide some 10 examples of areas where currently the limited access—by access, I mean reasonably assured access within budget affecting our material surface science work, both in the basic research area and also in the application to tribology.

Again, I would like to emphasize research on both the applied side and the basic side. For example, our fundemental understanding of surface science is many years ahead of important applications in tribology, theory and practice of lubrication and wear. We have capabilities to do extensive computer analysis of surface characteristics which, because of inadequate access to supercomputers, are not being utilized by the applied conservation programs.

We are just beginning the tribology program at the laboratory, building on the basic research side. But I think that is just indicative of the weakness in this system.

There are also excellent examples in the fission reactor physics and reactor safety areas in which we are employing simplified models to stay within the limits of routinely available machines rather than exploring the full complexity of the underlying theories. For example, indications are that there is a lot to be gained in enhanced understanding of the performance of reactor containment vessels.

Mrs. LLOYD. Do you also feel that computer resources are really being used to their full ability?

Dr. ROBERTS. My understanding is that the available DOE resources, for example, the Magnetic Fusion Computer Center, are being used to their fullest. They are saturated when there is just this tremendous latent demand. So what the researcher does is sculps his problem to what tools he has rather than to the state-of-the-art of the technology. It is as simple as that.

Mrs. LLOYD. Dr. Lester, would you like to comment?

Dr. LESTER. Yes, I would like to respond on this general point. From my experience in industry, consisting of over nine years at IBM Research in San Jose, Calif., I think there is an intrinsic aspect which one benefits from by having sizable amounts of resource available, which was the case for me there. And that is a research program can improve dramatically if one is not immediately resource limited.

A computationally oriented research program, regardless of the problem area and we can enumerate many of them and many have been enumerated before, changes radically if the limiting factor is not computer resources.

I think the kind of research, both in applied and basic areas that we will see if that limitation is removed will be significant and far-reaching.

I think this is a very important point and I want to emphasize it strongly.

Mrs. LLOYD. Pardon me. I would like to ask all of you this question. Do you feel the urgent aspects of the supercomputer situation and if it is true, what are the special time critical items that need to be addressed at this point? The urgent aspects that need our attention.

Dr. HAUSSMANN. An interesting little piece of mathematics—the Japanese expect in the next 6 years to have a thousandfold improvement over our leading edge computer now—if you do the mathematics on that and spread it out over time, it doesn't quite work that way, but it is roughly a factor of three in each of those years. They will not quite work it that way but that is what it turns out to be.

Then when you couple that in with the fact that their program has been underway a few years—to lessen the gap of the situation where they are ahead of us, we do have to get underway right away.

There are some things that we can do very rapidly that would be very advantageous, and providing current supercomputer time to organizations that have very little of it, namely, the universities and the energy research programs, would be an obvious first step. And the sooner you could do it, the better.

Also, getting underway on enhanced levels of development funds would be very good.

Mrs. LLOYD. Thank you. Would either of you gentlemen like to comment further?

Dr. LESTER. Yes, I would. I think there are two aspects to be considered. One is the long-range aspect of improved computing machines, more specifically, the support of fundamental research necessary for their development. I think it requires action which recognizes our needs at the present time, that is, improved support, greater support, for research on better machines.

The other dimension of the problem is one alluded to earlier, namely, the educational aspect, which requires access to current class VI type machines. I think the problem of building for the future is so critical, as also indicated by remarks made by the gentleman on my left, that we have to get going and take the necessary steps immediately. I think that is a very urgent requirement.

Mrs. LLOYD. Dr. Roberts?

Dr. ROBERTS. I can only add that I think there is a tremendous latent demand, reflected in the Argonne computing five-year plan, and I think the Oak Ridge 5-year plan. Both call for major machine acquisitions on the time frame of the next year or two. I think that is the way to, if you will, bootstrap the operation and assure the researcher the opportunity of access which moves that researcher into an entirely different planning framework, as Dr. Lester points out, and provide the funds for the purchase of machines at installations where they can fit into existing computer management organizations, such as at the national laboratories and certain major universities.

I think the sooner we get going in that immediate direction, the better.

Mrs. LLOYD. Thank you. I think I have used my time rather generously. Mr. Lowery?

Mr. LOWERY. Thank you, Madam Chairman.

If we were to pursue a policy as you all seem to be suggesting and there being a consensus of having supercomputers available for research primarily our academic institutions through the national laboratories, what are we looking at by way of cost? How many would be needed and how much access would there be, and would it be enough?

Dr. HAUSSMANN. I will make a couple of remarks on that. The National Magnetic Fusion Computer Center which provides state-of-the-art computer time technology to all of the magnetic fusion community is about $15 million a year and its major computers are a couple of Crays.

Mr. LOWERY. Are those operating costs or does that include the capital costs for the equipment itself?

Dr. HAUSSMANN. A lot of the costing—there was a change, rightly or wrongly, a few years ago where main frames used to be strictly a line item in Congress and in the last few years they have tended to be lease-purchase arrangements. So they tend to show up in operating costs, so the cost could be distributed over, say, 3 or 4 years.

So that really contains effectively much of the capital costs problem for all intents and purposes.

I think I disagree with the gentleman from Argonne. I think the antein fee on computer centers is appreciable enough that I would say the Magnetic Fusion Computer Center is at the lower edge of cost effectiveness if you doubled that, say, four leading edge computers instead of two, you would get an even better buy—I can't give you off the top of my head a number, but it certainly would be distinctly less than a factor of two increases in cost.

I personally believe that while you may want to sample stand alone computers in some unique circumstance, these distributed nets that Dr. Lester talked about also, are probably the way you are going to want to provide most machine time. It is not terribly important where that time is generated as long as it gets to the users.

Mr. LOWERY. Is there a maximum utilization of the existing computers or is there a capacity, surplus capacity?

Dr. HAUSSMANN. I can only speak for Livermore. We run them 24 hours a day, 365 days a year.

Mr. LOWERY. You are maxed out.

Dr. Lester?

Dr. LESTER. With regard to Lawrence Berkeley Laboratory, we do not have class VI machines. Our largest machine is the CDC-7600 which is quite old and we are fully utilizing that resource. We do have in our long range plan a recommendation for replacement equipment.

So we want state-of-the-art equipment there.

Mr. LOWERY. Your suggestion of using the national laboratories as regional centers, has that been costed out?

Dr. LESTER. I do not have firm figures on that.

Mr. LOWERY. Give me estimates?

Dr. LESTER. No, I am unable to provide them.

Mr. LOWERY. How many would——

Dr. ROBERTS. Let me, it was actually my suggestion of regional centers, but I think that the numbers of $15 million a year is represented by the Magnetic Fusion Energy facilities are comparable. I think that figuring on one or two machines and either capitalizing or through the lease process—it depends on how you do your budgeting, I realize that.

Mr. LOWERY. How many centers?

Dr. ROBERTS. I really don't think that I can put a number on it. I really think the DOE has to. Because what happens is the budgets and the planning all is tuned to what is available and what we have been told. I will give you an idea. When time was made available on the Magnetic Fusion Energy System, to basic researchers throughout the Department of Energy's national laboratories the time was oversubscribed to by a factor of 15. They are just in tremendous demand.

I think the challenge comes in how deciding when to stop satisfying the marginal demand. I do know of researchers who will use all the computing power we can make available. In fact, I think someone has said that the supercomputer is almost defined by that state-of-the-art which is one step less than the researcher really wants.

The point is that there is in some sense an insatiable demand but at the same time there are many good examples, excellent examples, of very important problems that are just not being addressed because of inadequate access.

So I would say that I could see it, you know, on the order, at least in the next 2 or 3 years of two or three such regional centers and as part of a learning curve on how to do this kind of thing.

Finally, I would emphasize the important regional nature of such centers. I do not subscribe to the idea that researchers throughout the country should sit at terminals and communicate with the west coast for all of their computing.

While in principle the interface with the computer is in many cases transparent, there will nevertheless be a significant migration of computer science talent and users to those centers where the machines exist. It is just an understandable tendency.

Finally, I think that there is not a significant return to scale beyond the level of two machines or three machines at one site.

Mr. Lowery. Let's say we had three machines. How much access would there be? How many people would they be available to?

Dr. Roberts. I really wouldn't know how to estimate that. Perhaps somebody else here could. The numbers are very tricky because the diversity of use is so great. For example, most of the computing in reactor physics at Argonne is done by one small group of people on behalf of a major diversified national reactor physics program.

Mr. Lowery. Well, instead of people entities, then, universities, labs.

Dr. Roberts. I would say on the order up to 50 major user institutions funded by DOE.

Dr. Haussmann. Our experience at Livermore in the weapons program and related—in the DOE programs at Livermore we have on the Livermore site computers for Cray-I's, they are the leading edge computers, and 1,200 people tap those with terminals. Generally they do fairly small problems. There probably are a few tens of people that run the really large numerical modeling problems.

Those problems run typically a few hours long. If you try to make one that runs a few days long, it just turns out to be an impractical way when you try to iterate over and over. So your typical large problem, which people do want to get around to doing, using the large-scale computers for, is almost automatically limited, for the most part, to a few hours. But there are large numbers of small users, too.

Mr. Lowery. Gentlemen, a final question, if I may.

Several of you have mentioned leasing as opposed to outright purchase. Are there advantages in terms of next generation of computers and models that are available upgrades through leasing as opposed to purchase?

Dr. Lester. Our experience at Lawrence Berkeley Laboratory, that would be a very worthwhile way to go primarily because the technology is evolving sufficiently rapidly that there would be difficulties very likely if one were to purchase in a given year and when that particular machine were to be, say, 5 to 7 years, and the technology has moved swiftly, then one has to again come up with major capital equipment to cover costs.

Leasing option, I think, is an interesting one because one can then move dated equipment out fairly easily and bring in replacement state-of-the-art equipment.

Dr. HAUSSMANN. That is the reason we moved to it too, as DOE has generally.

Mr. LOWERY. Thank you, Madam Chairman.

Mrs. LLOYD. Thank you.

Mr. Lewis?

Mr. LEWIS. Thank you, Madam Chairman.

Mr. Haussmann, to what extent has the supercomputer manufacturing industry been exposed to your S-1 development? Please describe their responses.

Dr. HAUSSMANN. Perhaps the most important single element of the Livermore effort has been the design codes that allow you to design these computers with a smaller number of people. There are advanced types of computer aided manufacturing—just a second.

There has been a distribution of that design program to 180 different U.S. sources. We have on occasion held symposia describing the characteristics of this machine. As I mentioned, the last one was about a few weeks ago, sponsored by the Navy.

It is a very good way for the industry and laboratories to work together.

Mr. LEWIS. What amount of discussion has taken place relative to their manufacturing and selling such a computer?

Dr. HAUSSMANN. We have had preliminary discussions with IBM to see if that is feasible. I think that by having perhaps manufacturing representatives at a laboratory like our own, they would have a comfortable enough communications mechanism that they could determine whether they would want to do that or not.

We would have no objection whatsoever. It hasn't progressed that far yet.

Mr. LEWIS. Thank you, Madam Chairman. Thank you, Mr. Haussmann.

Mrs. LLOYD. Thank you very much, Mr. Lewis.

I want to thank you gentlemen for participating in the hearings today and I am sure that members who could not be present would like to submit questions for you to respond to in writing and, without objection, they will be included in the record. Thank you.

Mr. Haussmann.

[The questions to be furnished follow:]

Dr. HAUSSMANN. I mentioned in my prepared statement that we provided Bob Cooper of DARPA and Al Trivelpiece of DOE some detailed examples of how the laboratory could help out. Those white papers we have copies of here if the staff or members would like them.

Mrs. LLOYD. We will accept them for inclusion in the record as well, Mr. Haussmann.

Dr. HAUSSMANN. Thank you.

Mrs. LLOYD. Thank you very much.

[The documents follow:]

Lawrence Livermore National Laboratory
PUBLIC AFFAIRS OFFICE
Anthony P. X. Bothwell
Michael Ross (SOURCE)
(415) 422-4599
Home: (415) 455-8507

EMBARGO FOR RELEASE
P.M.s Friday, May 20, 1983
NR 83-05-03

NEW LIVERMORE LAB LASER PROCESS 'PAINTS' INTEGRATED CIRCUITS ON SILICON

BALTIMORE -- Pulses of laser light flashing millions of times per second through selected gases have been used to "paint" integrated circuits directly onto silicon wafers, scientists from Lawrence Livermore National Laboratory (LLNL) in California announced today (Friday, May 20, 1983) at an international laser research conference here.

When fully developed, the revolutionary new process, called "laser pantography," and a Livermore-created computer code for making integrated circuits can reduce the time needed to design, make and test prototype and limited production specialized circuits from weeks to a few hours.

"We hope to have a system painting 1,000 transistors a second by the end of the year," said Dr. Irving Herman, a physicist in the LLNL Special Studies Group and co-leader of the research team with Physicist Bruce McWilliams.

It will be possible, the researchers said, to inscribe overnight the circuits of today's most powerful computers onto a silicon wafer 5 inches in diameter -- more than a 1,000-fold reduction in size. Personal computers of the future using a single silicon wafer could easily have the power of today's supercomputers costing about $10 million, the Livermore scientists said.

Though similar laser processes have been under study at a number of academic and industrial laboratories, the Livermore group claims to be the first to have produced operating circuits.

McWilliams described the new process at the Conference on Lasers and Electro-Optical Systems, meeting at the Baltimore Convention Center this week.

LLNL is operated for the United States Department of Energy (DOE) by the University of California. The staff of more than 7,300 at Livermore conducts research in nuclear weapon design and testing, nuclear fusion, basic physics, energy, biomedical and environmental sciences. LLNL researchers use one of the world's most powerful computer centers to model and predict the results from scientific experiments.

Current supercomputers can perform up to 60 million mathematical operations a second, about as fast as the entire human race can compute using pocket calculators and at least 10,000 times faster than personal computers, said Dr. Lowell Wood, head of the LLNL Special Studies Group.

"We need computers that are 100 to 1,000 times faster than those now available," said Carl Haussmann, LLNL Associate Director at Large. "We are working with the U.S. Navy and DOE to develop a new type of 'multi-processor' supercomputer, the S-1, and are very interested in U.S. government, industry and universities working together, as they are in Japan, to advance our country's capability to make and use ever more powerful supercomputers."

The new Livermore laser pantography process could greatly increase the number of supercomputers available for use, shorten the time to develop improved computers and dramatically reduce their cost.

"There are about 100 supercomputers in use in the world today," Haussmann said. "I could imagine 10 of these laser pantography units sitting in a single room with each making one supercomputer silicon wafer a day. That could mean several thousand new supercomputer modules could be built in that one room in a year."

Laser pantography can change the way integrated circuits are made

An integrated circuit is merely a tiny, thin piece of a perfect silicon crystal emblazoned with a stack of intricate patterns. Each section of the pattern has the special electrical properties needed to form the transistors, capacitors, resistors and connectors that form the complex electrical and logical network that comprises the brain of a computer.

The traditional method for making integrated circuits involves repeated series of processes such as photographic "burning" of the silicon surface, coating of thin layers of important electrical materials, and etching away unwanted areas. Especially for prototype circuits, the process is slow, expensive and, worst of all, yields a relatively low fraction of usuable circuits. Most are "born dead" -- irrevocably damaged during the laborious processing.

In contrast, the Livermore laser pantography approach shows potential to be fast, direct and reversible. If a design is changed, a new circuit can be made easily from scratch. If a mistake is made, it can be erased and corrected.

Laser pantography involves rapid reactions using intense green laser light directed onto silicon surfaces with an intensity a billion times greater than noontime sunlight. At times, various gases are introduced into a reaction chamber. The laser is pulsed on and off so fast that reactions occur only in the center of the focused laser spot, about one micron (a millionth of a meter or about one twenty-five-thousandth of an inch) in diameter. This equals the resolution of current integrated circuit technology.

The intense light superheats the silicon surface for such a short time that heating and cooling rates of hundreds of billions of degrees per second are achieved, giving the laser pantography process its great sensitivity and resolution. When the gases are admitted into the chamber, reactions occur on the hot surface that can remove or change the electrical properties of the silicon, or deposit the desired special electrical materials.

It's like 'a set of laser paint brushes'

McWilliams, an amateur painter in his free time, describes laser pantography as "a set of laser paint brushes, which we dip into paint-pots of various exotic gases in order to paint on a canvas of silicon."

The resulting "paintings," however, become the brains of high-powered computers when electricity is applied.

A major thrust of the Livermore effort is to couple the laser pantography process with both a computer laser controller and another computer program that designs and performs the placement of the circuit elements on the chip. A totally automated system even holds the promise that powerful computer systems can "reproduce" themselves without the extensive human participation and effort now required.

"New generations of ever more powerful computers can thus evolve in days rather than in the years that currently separate generations," said Dr. Lowell Wood, head of the LLNL Special Studies Group.

Co-authoring the Baltimore paper with McWilliams were Irving Herman, Rod Hyde, Fred Mitlitsky and Wood, all members of the Livermore Laboratory's Special Studies Group. Mitlitsky is also an electrical engineering graduate student at Columbia University. His Columbia advisor, Prof. Richard Osgood, is a consultant to the Livermore effort. Collaborating on this research were Hon Wah Chin, Andrew Weisberg, Elon Ormsby, Rich Mendonca and Don Duerre.

Lawrence Livermore National Laboratory

BLIC AFFAIRS OFFICE
thony P. X. Bothwell
Michael Ross (SOURCE)
(415) 422-4599
(Home) (415) 455-8507

FOR IMMEDIATE RELEASE
P.M.'s Friday, May 20, 1983
NR 83-05-03 (Technical)

LASER PANTOGRAPHY-A BREAKTHROUGH IN COMPUTER CREATION WITH LASERS

Creation of integrated circuits directly by focusing pulsed laser beams onto semiconductor sheets was announced today by scientists from the Lawrence Livermore National Laboratory. Such circuits, which are used as the active portions of both personal computers and the largest computers in existence, have never before been directly created via any means, though scientists have been striving toward this particular goal for several years.

Speaking to researchers gathered from all over the world at the Conference on Lasers and Electrooptical Systems now meeting in Baltimore, Dr. Bruce McWilliams, leader of the Livermore Lab team that did the work, related both the ground-breaking experimental results in what he called "laser pantography" and the theoretical advances that led up to them.

McWilliams noted that all integrated circuits have been created hitherto by a long series of exacting processes, consisting of writing one pattern after another onto the surfaces of semiconductors such as silicon and then "burning" these patterns permanently into the silicon surface. Typically, he said, this patterning is done by photographic means, imaging a master pattern onto an analog of photographic film on the surface. This exposed "film" is then developed, and the image impressed into the surface, often by high temperature processing. When the appropriate patterns are stacked on top of each other in the right sequence, then transistors, resistors, capacitors and interconnecting wiring--the building blocks of modern microelectronics--are combined to create an integrated circuit. However, even the best modern integrated circuit manufacturing processes are slow, expensive, and prone to irreversible errors.

Industrial technicians today typically move the semiconductor wafers--thin circular sheets 3 to 5 inches in diameter--from a patterning station in and out of one or more chemical reaction vessels, and then cycling this process through a dozen or more patterning steps over a period of several days. In the end, perhaps 90% of the integrated circuits are "born dead" --inadvertently damaged beyond repair in the creation process.

In the completely different "direct" approach pioneered by the Livermore Lab team, the wafer is loaded into a small chamber with a transparent side, where it remains throughout the processing. A focussed green laser beam, typically switched on and off millions of times per second, is then played over the surface of the wafer under the control of a computer. In some stages of the processing, a special gas mixture is introduced in the chamber before the focussed laser beam is brought to bear, while before other stages, the chamber is pumped down to a very good vacuum.

The laser beam is typically focussed to a spot size less than a ten-thousandth of an inch across, to a diameter only a few wavelengths of light, and to an intensity a billion times greater than sunlight at noon. Under these conditions, the silicon surface is heated from room temperature to its melting point in a few billionths of a second and, because the spot size is so small, cools down comparably rapidly when the laser beam is turned off or moved to another spot. Heating and cooling rates of hundreds of billions of degrees per second are thereby realized.

The laser-superheated surface reacts violently, but very briefly, with selected gases which have been admitted into the chamber in which it is contained. Some of the gases etch the tiny heated spots on the surface, removing whatever material may be present at rates thousands of times greater than are typical of material removal by milling machines and lathes in machine shops. Other gases react with the microscopic hot spots on the surface to deposit solid material, rather than etch it away, also at very high rates. Still third types of gases react with the laser-heated microscopic spot on the semiconductor surfaces to merely change its composition, and thus its electrical properties.

"While it may seem impossible to do anything useful in a millionth of a second or less while any given spot on the surface is laser-heated and reacting with the gas right above it, you have to remember that we may be removing or adding only a trillionth of a gram of material or a thousandth of a trillionth of a gram of dopant in order to create a one micron-scale transistor," noted Dr. Irving Herman, a senior Livermore Lab scientist working on this project. Sometimes, for instance when a precisely contaminated surface is transiently melted in vacuum by the laser beam, the goal is merely to move a billion atoms of the contaminant a few millionths of an inch in position, in order to create a certain type of transistor structure.

"The technology which we've created," remarked McWilliams, an amateur painter in his free time, "is a set of laser paint brushes, which we dip into paint-pots of various exotic gases and then paint with on a canvas of silicon. The striking thing is that the resulting "paintings," in addition to being beautiful to look at, can be created literally at the speed of light and then will compute for us at very high speeds when we apply electricity to them, all in the same container." He went on to explain that not only is the motion of the focussed laser beam over the wafer surface computer-controlled, but that a small, dedicated computer also turns the beam on and off, in addition to varying its turned-on intensity to "dial" the temperature of the surface when the beam is focussed onto it. The pressure and composition of the gas mixture present in the wafer-holding chamber at any time can also be computer-prescribed.

"Thus we can tell the computer what kind of circuit we want created on any given part of the wafer, start it up, and then go out to dinner while it writes the corresponding patterns for us from its detailed knowledge of how to go from a specified circuit to the required set of stacked patterns on the wafer surface," McWilliams said. "We hope to have it directly writing as many as a thousand transistors per second later this year, so that it will then be able to write the complete set of microelectronics of the largest computer in existence literally overnight. If it makes a mistake or we change our minds about what we want, it can erase whatever portion of the patterning should no longer be there, and re-write what is wanted, just like erasing a portion of a blackboard and re-writing it."

Two of the Livermore Laboratory's leaders hailed the work. "I wanted them to use this splendid new technology to write integrated circuitry over a square meter of semiconductor surface," declared Edward Teller, the prize-winning nuclear physicist and technologist, "but they convinced me that nearly everything worth doing could be one in one-hundredth of this area. It's difficult to overestimate the ultimate significance of this breakthough." Carl Haussmann, Associate Director-At-Large of the Laboratory, observed, "With the technology already in hand, they could potentially put all of the high-speed memory capacity of one of the world's most powerful computers onto a single silicon wafer, without human intervention. The long-term implications are mind-boggling." In a laser pantography paper presented last fall, the Livermore researchers noted that their new technology brings much closer the day when powerful computer systems can "reproduce" without the extensive human participation and long effort periods which are currently required.

Co-authoring the paper with McWilliams and Herman were Rod Hyde, Fred Mitlitsky and Lowell Wood, also members of the Livermore Lab's Special Studies Group. Mitlitsky is also a graduate student in Columbia University's Electrical Engineering Department, and his Columbia adviser, Professor Richard Osgood, is a consultant to the Livermore Lab effort. Hon Wah Chin, Andrew Weisberg, Elon Ormsby, Rich Mendonca and Don Duerre also collaborated in the work reported.

The Livermore Laboratory is operated by the University of California for the U.S. Department of Energy, under whose auspices the reported work was performed.

/js
1482a

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government thereof, and shall not be used for advertising or product endorsement purposes.

WAFER-SCALE LASER PANTOGRAPHY:
I. PYROLYTIC DEPOSITION OF METAL MICROSTRUCTURES

IRVING P. HERMAN, RODERICK A. HYDE, BRUCE M. MCWILLIAMS,
ANDREW H. WEISBERG and LOWELL L. WOOD.

Physics Department
Lawrence Livermore National Laboratory,
Livermore, CA 94550

ABSTRACT

Mechanisms for laser-driven pyrolytic deposition of micron-scale metal structures on crystalline silicon have been studied. Models have been developed to predict temporal and spatial properties of laser-induced pyrolytic deposition processes. An argon ion laser-based apparatus has been used to deposit metal by pyrolytic decomposition of metal alkyl and carbonyl compounds, in order to evaluate the models. These results of these studies are discussed, along with their implications for the high-speed creation of micron-scale metal structures in ULSI systems.

INTRODUCTION

In order that state-of-the-art computers can reproduce with less human interference and thus with lower error rates and latencies, it is necessary for them to be able to micropattern semiconductor substrates directly and at high speed. Contemporary supercomputer systems composed of 10^8 transistors, each constituted of as many as 10^2 'tiles' of various materials stacked in particular arrangements, contain as many as 10^{10} 'tiles' in their ultra-large scale integrated (ULSI) patterns. In order that latencies involved in computer-directed generation of new computer-scale micropatterns be not much greater than a day, creation of a single 'tile' must be accomplished in an interval of the order of 10^{-5} seconds.

In the series of papers of which this is the first, we will report on means for accomplishing such processing in this context. This paper reports on that aspect of laser micropatterning which is possibly the least advanced through the present: the discretionary, high-speed generation of lines of metal—paths of metallic 'tiles' laid edge-to-edge—suitable for interconnecting active devices in ULSI circuits.

---

* Work performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under contract number W-7405-ENG-48.

Monolithically implemented computers of the present era consist of sheets of metal traces of micron-scale width and thickness, separated almost (but not quite) everywhere from each other by nearly hole-free dielectric sheets, and occasionally touching the terminals of transistor devices which are embedded for convenience in a semiconductor substrate common to all such devices. The large majority of the structure (indexed by the fraction of total 'tiles' in the pattern) of computers implemented in the ever more popular metal-oxide semiconductor (MOS) family of technologies is expressed in metal (or in surrogate metal, highly doped polysilicon). It is thus quite crucial to greatly extend the hitherto highly limited ability to create suitable metal microstructures directly with laser processing techniques, as this approach is qualitatively superior to all other current prospects for high-speed, computer-controlled micropattern generation directly on semiconductor substrates.

We report here an order of magnitude enhancement in the rate of metal micropatterning in the ULSI context, relative to the highest rates previously reported. Taken together with results from other workers on rapid semiconductor substrate doping and dielectric layer creation, this advance makes feasible in principle the generation of monolithic, ULSI supercomputers on day time scales.

The laser-driven metal deposition processes of interest consist of a semiconductor substrate surfaced with the reactant gas, with a laser beam focused to the region on the substrate where metal is to be deposited. Two different types of laser-induced metal deposition processes are candidates for this role: pyrolysis and photolysis. The pyrolytic processes are based on thermal decomposition reactions which occur due to laser-induced heating of the substrate, while the photolytic processes entail gas phase photolytic decomposition reactions occurring in the laser-irradiated gas-filled region above the substrate.

The ultimate limit on the rate of metal deposition by either pyrolytic or photolytic processes is determined by the maximum possible rate of mass transport of the reactant gas to the substrate. Pyrolytic processes may usefully be thought of as unimolecular decompositions, while photolytic ones are two-body (binary) ones in the same sense—both gas molecules and photons must occupy the same space-time volume for photolysis to occur.

In order for photolysis to occur in unit time, the photon fluence must equal the reciprocal of the cross-section for the metal-bearing molecule to photolytically absorb the photon. Since the unit of time is that for the molecule to transit the reaction zone—about 10 nanoseconds for molecular weights of the order of 100 amu and reaction zone scales of the order of 1 micron—and since the largest photolytic cross-sections of interest less than 10^{-17} cm^2, the photon fluxes required for mass transport-limited photolysis are of the order of 10^{25} photons cm^{-2} sec^{-1}, or about 10 megawatt cm^{-2}, which implies a minimum input of 0.1 watts of hard-won 4-6 eV laser photons.

By way of comparison, the laser intensity required for pyrolysis of metal-bearing gases of interest at a comparable mass-processing rate is just that needed to maintain a square micron of substrate at a temperature of the order of 1000°K, which is again about 10^7 watts cm^{-2}, but now involves an input of 0.1 watts of readily available, wavelength-insensitive optical laser spectral power. Thus, even when using metal-bearing gases with the highest UV photolytic cross-sections, the laser intensity regime for efficient photolysis already completely overlaps the one for pyrolysis, but is accessed with substantially greater technical difficulty. When using gases with more typical photolytic cross-sections, the laser intensities required for efficient photolysis will boil the substrate on nanosecond time scales. Intensity reduction to more acceptable levels will slow the photolysis rate, relative to the pyrolytic one, by the ratio of the photolytic cross-section to the peak value of 10^{-17} cm^2, e.g., for aluminum trimethyl with its photolytic cross-section of 10^{-20} $cm^{-2}(\lambda \simeq 0.26\ \mu m)$, photolytic deposition of aluminum metal will be approximately three orders of magnitude slower than will pyrolytic.

We therefore restrict our attention to pyrolytic mechanisms of metal micropatterning of semiconductor substrates. Earlier work has demonstrated photolytic deposition of Al, Cd, Zn and Sn from gas-phase and surface decomposition of the respective alkyls[R1], and of Fe, W, Cr, and Mo from the respective carbonyls[R2,R3], using ultraviolet ($\lambda \leq 0.26\ \mu m$) laser light. Previous workers have pyrolytically deposited large, dendrite-rich strips of Al, Zn and Cd at low speeds on GaAs surfaces with focussed radiation from a krypton-ion laser[R4], and deposited large scale Ni[R5] and W[R6] films with CO_2 laser radiation focused on silica surfaces.

Mass transport-limited, laser-driven pyrolysis of dimethyl zinc, trimethyl aluminum and nickel carbonyl to form metal traces with micron-scale widths and thicknesses on silicon wafers by focused argon-ion laser light ($\lambda = 0.5145\ \mu m$) is reported in the present paper. Because of its superior properties—high $Ni(CO)_4$ vapor pressure, high melting point, high electrical conductivity and excellent corrosion resistance of metallic nickel, and good model qualities for other systems of interest—this nickel system has been singled out for extensive study. We interpret these results with theories of quantitative predictive quality, and thereby extrapolate them to ULSI applications of ultimate interest.

THEORY OF LASER-INDUCED PYROLYTIC DEPOSITION PROCESSES

The laser-induced pyrolytic metal deposition processes under consideration are based on thermal decomposition reactions of the form:

$$[MX_n] \rightarrow [M] + n[X] \tag{1}$$

where MX_n is the reactant gas molecule (e.g., a metal alkyl or carbonyl), M is the metal atom being deposited, and X is a product gas of the reaction. The process can be thought of as involving five consecutive steps:

a. Diffusion of the reactant gas to the heated portion of the substrate;
b. Adsorption of reactant;
c. Decomposition of the reactant;
d. Desorption of the volatile product; and
e. Diffusion of the product gas away from the substrate.

At lower temperatures, the rate-determining step of the process is usually (c); at higher temperatures, where the decomposition reaction rate becomes substantial, (a) and (e) become the rate-limiting steps.

We present a generally applicable model for high-speed deposition of metal microstructures by pyrolysis of a metal-bearing gas, emphasizing generation of micron-scale nickel traces by pyrolysis of $Ni(CO)_4$.

The kinetics of processes (b) through (d) have been analyzed by Carlton and Oxley[R7]. Their expression for the overall reaction rate is:

$$R(p_1, p_2, T) = R_0(T) f(p_1, p_2, T) \tag{2}$$

where

$$R_0(T) = 5.1 \times 10^{29} exp\{-T_0/T\} \quad cm^{-2} sec^{-1}; \quad T_0 = 11,022^\circ K \tag{3}$$

$$f(p_1, p_2, T) = k_1^2 [p_1^2 - (p_2^4/K_{eq})^2] / [1 + k_1 p_1 + k_2 p_2]^2 \tag{4}$$

$$\begin{aligned} k_1 &= 0.089 exp\{T_1/T\} \quad Torr^{-1}; \quad T_1 = 352^\circ K \\ k_2 &= 0.00152 exp\{T_2/T\} \quad Torr^{-1}; \quad T_2 = 2,013^\circ K \\ K_{eq} &= 1.4 \times 10^{31} exp\{-T_3/T\} \quad Torr^3; \quad T_3 = 18,800^\circ K \quad [R8] \end{aligned} \tag{5}$$

Here p_1 is the $Ni(CO)_4$ partial pressure, while p_2 is that of CO. The reaction is second order and is choked at high pressure by adsorbed $Ni(CO)_4$ and CO. Since f is always less than one, R_0 represents the peak rate possible, which is attained at high p_1 and low p_2. We can neglect the reverse reaction, since at the temperatures needed to make R_0 large, the equilibrium constant for the reaction, K_{eq}, is also large; the recombination reaction is thus negligible for present purposes.

Examination of (2) shows that, as we increase surface temperature, the rate will climb due to the activation energy factor R_0. Eventually, though, we will reach an asymptotic condition where gas diffusion is insufficient either to maintain p_1 near its far-field value, or to disperse CO and keep p_2 low near the deposition site. Both of these effects will cause f to fall. Our model must therefore relate T to laser power, and must consider gas diffusive mass transport. First we obtain $R_0(T)$.

The temperature field produced in a plane slab of solid silicon by a linearly swept laser beam of Gaussian profile has been treated previously[R9, R10]. Using a Kirchhoff transform we define a linearized temperature:

$$\theta(T) = \int_{T_r}^{T} \frac{K(T')}{K(T_r)} dT', \tag{6}$$

where T_r is the reference (initial) temperature. Writing the space- and time-dependent laser intensity I as

$$I(x - vt, y) = \frac{P}{\pi r_0^2} exp\left\{-\frac{(x - vt)^2 + y^2}{r_0^2}\right\} \tag{7}$$

where P is the incident laser power and v s the laser focal spot sweep rate, then

$$\theta(X, Y, Z; V) = \frac{2}{\pi}\theta_c \int_0^\infty exp\left\{-\frac{(X + Vu^2)^2 + Y^2}{1 + u^2} - \frac{Z^2}{u^2}\right\}\frac{du}{1 + u^2} \tag{8}$$

where

$$X = \frac{x}{r_0} \qquad Y = \frac{y}{r_0} \qquad Z = \frac{z}{r_0} \qquad V = \frac{vr_0}{\sqrt{8D}} \tag{9}$$

$$\theta_c = \frac{P(1 - \Re_S)}{2\sqrt{\pi} K(T_r) r_0} \tag{10}$$

This temperature solution is found via a Green's function route. This approach is not exactly correct, since the silicon thermal diffusivity D is temperature-dependent, but for "relatively slow" scan velocities, D fades from significance.

The reflectivity $\Re_S$ has been assumed constant over the laser beam focal spot, which is reasonable for solid Si. The term θ_c is the peak value of θ for the static case; to relate it to the actual temperature, we must assume an expression for thermal conductivity. We use $K(T)$ and $\Re_S$ from[R9]:

$$K(T) = \frac{k}{T - T_k}; \qquad k = 299\ W/cm; \qquad T_k = 99^\circ K \tag{11}$$

$$\Re_S(T) = 0.367 + 4.29 \times 10^{-5} T$$

Solving (6) gives

$$\theta(T) = (T_r - T_k) ln \frac{T - T_k}{T_r - T_k}. \tag{12}$$

$$T(\theta) = T_k + (T_r - T_k)exp\left\{\frac{\theta}{T_r - T_k}\right\} \tag{13}$$

$$T_c = T_k + (T_r - T_k)exp\left\{\frac{P(1 - \Re_S)}{2\sqrt{\pi}kr_0}\right\} \tag{14}$$

For $P = 300$ mW and $r_0 = 2\,\mu m$, we find $T_c = 575°$K.

The simplest expression for metal deposition rate is obtained by assuming a relatively slow laser focal spot scan rate and looking at the center of the spot; placing (14) in rate (3) we obtain the rate of increase of the metal film thickness $\dot{W}_c$:

$$\dot{W}_c = 5.6 \times 10^{10} exp\{-T_0/T_c\} \qquad \mu m/sec \tag{15}$$

This is plotted in Figure 1 as a function of specific laser power P/r_0, using $T_r = 300°K$. This model can also predict the spatial variation of growth rate by combining (3), (8) and (13). We can obtain a simple relation by Taylor expansion, which is useful since $T << T_0$

$$\theta(r) \approx \theta_c[1 - \frac{r^2}{2r_0^2}] \tag{16}$$

$$T(r) \approx T_c - \theta_c \frac{T_c - T_k}{T_r - T_k} \frac{r^2}{2r_0^2} \tag{17}$$

so

$$\dot{W} \approx \dot{W}_c exp\{-r^2/\rho^2\} \tag{18}$$

where ρ is the effective interaction radius

$$\frac{\rho}{r_0} = \frac{T_c}{\sqrt{T_o \theta_c}} \sqrt{2 \frac{T_r - T_k}{T_c - T_k}} \tag{19}$$

For example, at $P/r_0 = 0.15\ W/\mu m$, we find $\rho/r_0 = 0.38$, so the Ni spot size should be about 2.5 times smaller than that of the laser.

The growth rates of Figure 1, based on Equation 15, are invalid at high temperatures, because gas diffusion cannot supply $Ni(CO)_4$ or remove CO sufficiently rapidly. Therefore, we need to obtain $f(p_1, p_2, T)$ and do so by analyzing gas diffusion by a Green's function approach analogous to (8). Solving for the "relatively slow" scan case using a reaction rate spatial variation as in (18), the species population changes at spot center are

$$\Delta n_i = \frac{\sqrt{\pi}}{2} \rho \frac{R_i}{D_i} \quad cm^{-3}; \quad i = 1, 2 \tag{20}$$

Here the rate for $Ni(CO)_4$, R_1, is $-R$ from (2), while R_2 for CO is $+4R$. The gas diffusivities are D_1 and D_2, while Δn_1 and Δn_2 are the Ni and CO population changes from their far-field values. If we convert populations to pressures, we arrive at equations which must be solved self-consistently:

$$p_1 = p_{1\infty} - 5.35 \times 10^{10} exp\{-T_0/T_c\} \frac{\rho T_r}{D_1} f(p_1, p_2, T_c) \qquad Torr \tag{21}$$

$$p_2 = 4 \frac{D_1}{D_2} (p_{1\infty} - p_1) \qquad Torr \tag{22}$$

where $p_{1\infty}$ is the pressure far from the laser beam focal spot in Torr, and other variables have units as noted above.

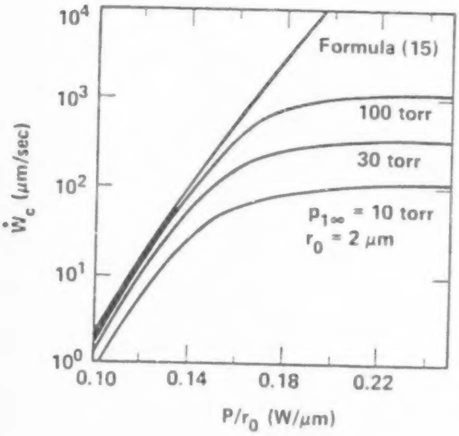


Figure 1 - Calculated Ni growth rate on Si substrate. The three right-most curves assume $r_0 = 2$ μm.

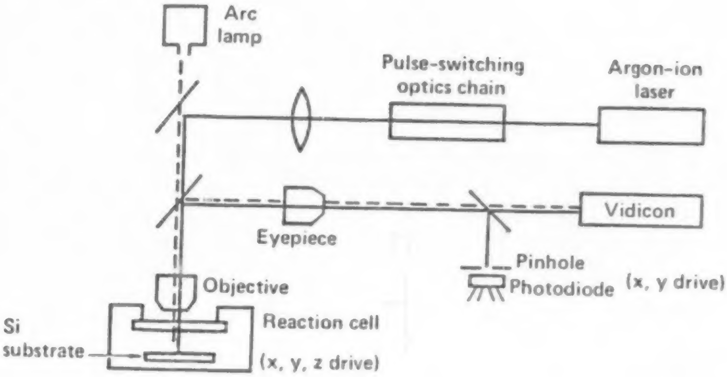

Arc lamp
Pulse-switching optics chain
Argon-ion laser
Vidicon
Eyepiece
Pinhole
Photodiode
(x, y drive)
Objective
Si substrate
Reaction cell
(x, y, z drive)

Figure 2 - Experimental arrangement for laser-pyrolytic deposition of metal microstructures.

The effects of diffusion are illustrated in Figure 1 for a 2 μm radius laser spot. We assume that the gas was at $T_r = 300°K$, not at T_c. For D_1 and D_2, we assume a 1 atm He buffer gas background, a common choice in our early experiments. Using the method of[R11], we obtain

$$D_1 = 0.23\ cm^2/sec; \qquad D_2 = 0.93\ cm^2/sec \tag{23}$$

The inaccuracy in using (15) at low laser beam power levels is due to neglecting $f(p_{1\infty}, 0, T_c)$. Of more fundamental interest are the high power asymptotic results. This diffusion-limited rate is

$$R \lesssim \frac{2}{\sqrt{\pi}} \frac{D_1}{\rho} n_{1\infty} \tag{24}$$

Expressed in terms of metal-bearing gas pressure and growth rate of metal thickness, it is

$$\dot{W}_c \lesssim 9.1 \times 10^{-4} p_{1\infty}/\rho \quad \mu m/sec. \tag{25}$$

where $p_{1\infty}$ is in Torr and ρ is in cm. This limit is set by gas transport to the laser-heated spot on the surface; other factors enter only weakly through the radius of the metal spot, ρ. Small spot sizes and high partial pressures of $Ni(CO)_4$ are desirable in order to realize high metal trace growth rates. Fortunately, this is the regime of most interest to us.

This analysis assumes that the laser spot is being scanned "relatively slowly" over the substrate—$v \lesssim 500$ cm/sec—and that deposition is taking place onto Si. Obviously, most of the Ni metal deposition occurs on top of existing Ni. Since only a thin metallic layer ($\lesssim 0.01\ \mu m$) is necessary to dominate reflectivity and absorption of laser light, we can assume that the entire laser spot sees the Ni reflectivity, $\Re_N$. Making the reasonable assumption that heat absorbed by the Ni transports primarily downward with little lateral conduction, then our analysis of the local Si temperature remains applicable for thin Ni layers, save for using $\Re_N$ in place of $\Re_S$. Using $\Re_N = 0.62$, the abscissa variable in Figure 1 now becomes $0.63P/r_0$. But our T_c is at the Si surface; the Ni surface is at a higher temperature and thus conducts the incident power down into the Si substrate. We can obtain a thickness limitation on the deposited metal film at which point the Ni starts to melt on its upper, laser-irradiated surface:

$$W_M = \frac{\pi r_0^2 K_N}{P(1-\Re_N)}(T_{mp} - T_c) \qquad cm \tag{26}$$

Here T_{mp} is the Ni melting point of $1728°K$, K_N is its thermal conductivity of $0.7\ W/cm/°K$, and T_c is found from (14), using a value of $\Re_N$ of 0.62. As an example, consider a r_0 laser spot radius of $2\ \mu m$ with a 0.6 watt incident laser beam power. We find that W_M in this case is $4.0\ \mu m$. This deposit thickness is attained quickly: a side effect of the large thermal gradient in the Ni is to drive the deposition process to the diffusion growth limit, so using (25) and (26),we find that melting occurs in a time t_M

$$t_M = 6.4 \times 10^7 (\frac{\rho}{r_0}) r_0^3 \frac{T_{mp} - T_c}{P p_{1\infty}} \qquad sec \tag{27}$$

For 10 Torr $Ni(CO)_4$, $t_M = 34$ msec for the conditions just stated.

Dwelling at a location for longer than this time will cause melting of the top surface of the Ni. If the molten Ni spreads out, the line width of the metal trace being laid down will grow; if it doesn't have time to spread then, as more Ni is deposited, further temperature rise occurs, which can culminate in vaporization and spattering of superheated liquid Ni.

As a result, a single pass *minimum* scan rate V_{min} which just avoids melting under the stated conditions is

$$V_{min} \gtrsim 2\rho/t_M = 3.1 \times 10^{-8} \frac{P p_{1\infty}}{r_0^2 (T_{mp} - T_c)} \ cm/sec \tag{28}$$

EXPERIMENTAL STUDY

Description of Apparatus

Polished p-type silicon targets, typically oriented in [100] with 30 Ω-cm conductivity, were suitably cleaned and placed in a reaction cell having a sapphire entrance window. The cell was filled with the desired partial pressures of buffer gas and metal alkyl or carbonyl, and then irradiated by an argon-ion laser ($\lambda = 0.5145\ \mu m$) focused by one of two microscope objectives (10X, NA = 0.25; 36X, NA = 0.50).

Desired metal patterns (frequently lines) were created by moving the reaction cell with an x-y table stepper drive (1 μm steps; 1500 $\mu m/sec$ maximum speed), and by amplitude modulating the argon-ion laser beam by an optics chain which often included a half-wave plate followed by an electro-optics switch that was sandwiched between two Glan-Thompson prisms. During the course of each run, the laser patterns on the substrate and the deposited metal features were viewed through a microscope via a vidicon outputting to a color monitor. Laser beam profiles on the substrate were measured using a pinhole-photodiode arrangement mounted on a calibrated x-y drive located after the microscope. The samples were analyzed using optical and scanning electron microscopy (the latter with X-ray fluorescence capability) and stylus profilometry.

Results

Operating with present experimental conditions of $\sim$ 10 Torr of metal alkyl or carbonyl and $\sim$ 0.4 W of 0.5145 μm radiation focused to $r_0 = 2 - 3\ \mu m$ on the Si surface, we observed rapid and localized deposition of Zn, Al and Ni from $Zn(CH_3)_2$, $Al(CH_3)_3$ and $Ni(CO)_4$. For scanning rates of $\sim 400\ \mu m/sec$, $2 - 6\ \mu m$ wide and $0.2 - 10\ \mu m$ high metallic lines were obtained in $\sim$ 4 sweeps. The aluminum lines possessed a $\sim 0.5 - 2\ \mu m$ diameter puffy, quasi-periodic substructure, in part due to the discrete stepping translation of the wafer, while the nickel substructure consisted of close-packed quasi-spherical beads of $\sim 0.5 - 4\ \mu m$ diameter. At high laser power levels, the line deposited of each metal contained a smooth-surfaced central axial depression due to melting and compaction of the beads within the depression, as discussed above. As was particularly evident in the Ni and Zn studies, there was significant vaporization of metal from this central valley and consequential prenucleation of the nearby substrate surface when high laser powers were employed.

Figure 3 demonstrates that the thickness of the deposited Ni strip is linear with the number of sweeps and with the dwell time per 1 μm stepper jump, as predicted by our theory. Figure 4 shows preliminary data concerning the height of the deposited Ni line versus laser power applied to the target area. These runs were performed with a gas mixture consisting of 11 Torr $Ni(CO)_4$ plus 700 Torr He. These data indicate that Ni was deposited at a rate of $\sim 100\ \mu m/sec$ at higher laser power, in agreement with the theoretical prediction of Eq. (25) and Figure 1.

Variations in height within a given metal line and between different lines drawn under similar conditions are attributed to the nature of nucleation on the surface. At lower power levels, the method of surface preparation (various methods of cleaning polished Si surfaces covered by its native oxide, including cleaning and stripping away the SiO_2) greatly affected the production of nucleation sites. Prenucleation due to boiling of metal from nearby nickel lines was observed to greatly lower the nucleation threshold for subsequent low laser power pyrolytic metal deposition runs. Nucleation effects are expected to be of much less importance in planned work involving substrate temperatures raised to far above threshold temperatures for sub-microsecond periods.

Sweeping (or otherwise time-modulating) a micron-scale laser beam spot relative to a location being metallized on a time scale of the order of 10 nanoseconds is predicted to be sufficient to intrinsically confine the deposition to within a 1 micron spot diameter on a silicon substrate, moreover very crisply, due to the exponential dependence of the deposition rate with surface temperature over the range of interest. Substrates of lower thermal diffusivity (*e.g.*, silica) are tolerant of more prolonged exposure before 'thermal transport blurring' of the laser beam spot will ensue.

Our present experimental results include creation of micron-scale metal lines at deposition rates of 0.1 millimeter thickness per second with argon-ion laser beam powers in the range 0.3-0.6 W incident on the silicon substrate with beam spot sweep rates of more than 500 microns per second.

At multi-watt incident laser beam power levels and saturated carbonyl pressure above the substrate, our experimentally validated theory predicts that nickel/iron lines with micron-scale widths and thicknesses of several tenths of a micron can be deposited on a substrate of choice at a sweep rate of 10 cm/sec. This is quite adequate for computer-driven creation of ULSI circuits at interestingly high rates.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to Dr. Richard Osgood for useful discussions during the course of this work, to James Atkins, Donald Duerre, Mark Garrett, Fred Mitlitsky and Elon Ormsby for their expert technical assistance in the experimental aspects of this study, and to Christine Ghinazzi for assistance in preparing this paper for publication.

REFERENCES

[R1] D. J. Ehrlich, R. M. Osgood, Jr. and T. F. Deutsch, IEEE J. Quant. Electron. **QE-16**, 1233 (1980).

[R2] D. J. Ehrlich, R. M. Osgood, Jr. and T. F. Deutsch, J. Electrochem. Soc. **128**, 2039 (1981).

[R3] R. Solanki, P. K. Boyer, J. E. Mahan and G. J. Collins, Appl. Phys. Lett. **38**, 572 (1981).

[R4] Y. Rytz-Froidevaux, R. P. Salathe and H. H. Gilgen, Phys. Lett. **84A**, 216 (1981).

[R5] S. D. Allen, J. Appl. Phys. **52**, 6501 (1981).

[R6] R. S. Berg and D. M. Mattox, "Proceedings of the Fourth International Conference on Chemical Vapor Deposition", edited by F. A. Glaski (Am. Nucl. Soc., Hindsdale, Illinois, 1973), p. 196.

[R7] H. E. Carlton and J. H. Oxley, A.I.C.L.E. Journal **12**, 86 (1967).

[R8] W. M. Goldberger and D. F. Othmer, IECPDE **2**, 202 (1963).

[R9] J. E. Moody and R. H. Hendel, J. Appl. Phys., **53**, 4364 (1982).

[R10] Y. I. Nissim, A. Lietoila, R. B. Gold and J. F. Gibbons, J. Appl. Phys., **51**, 274 (1980).

[R11] J. O. Hirschfelder, C. F. Curtiss and R. B. Bird, *Molecular Theory of Gases and Liquids* (John Wiley and Sons, New York, 1954) p. 14.

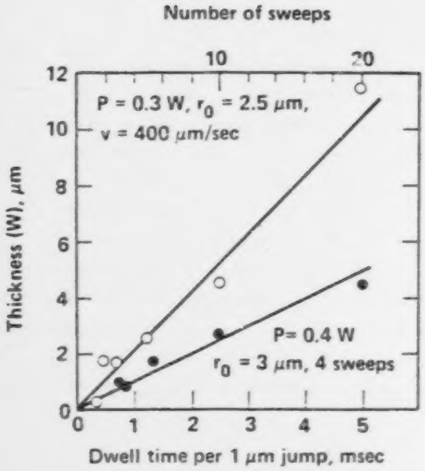


Figure 3 - Temporal dependence of nickel line deposition on silicon by laser pyrolysis (see text).

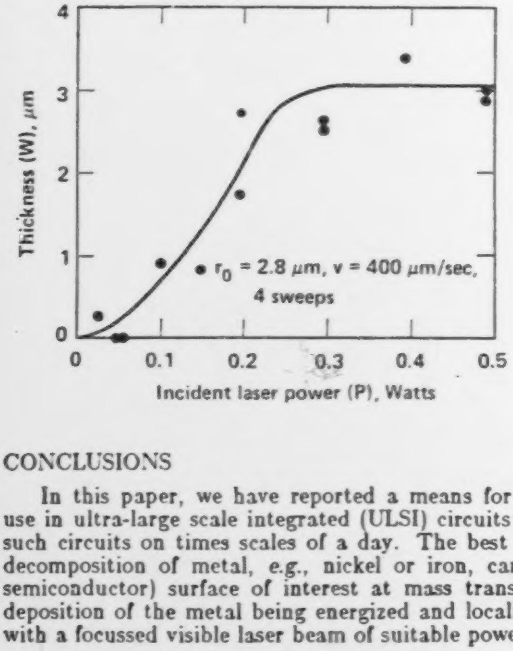


Figure 4 - Laser power dependence of pyrolytic deposition of nickel lines on silicon substrate (see text).

CONCLUSIONS

In this paper, we have reported a means for generating lines of metal suitable for use in ultra-large scale integrated (ULSI) circuits at rates which correspond to creating such circuits on times scales of a day. The best example of this approach involves the decomposition of metal, e.g., nickel or iron, carbonyl positioned above the (typically semiconductor) surface of interest at mass transport-limited rates, with the pyrolytic deposition of the metal being energized and localized by transiently heating the surface with a focussed visible laser beam of suitable power level.

The first-principles theory which we have presented is quantitatively predictive of the preliminary experimental results which we have obtained with respect to all measured aspects of metal deposition. It asserts that metal depth build-up rates of the order of centimeters per second can be attained with the use of nickel and iron carbonyls, with somewhat lower rates being attainable for the carbonyls of other metals of interest, such as molybdenum and tungsten.

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government thereof, and shall not be used for advertising or product endorsement purposes.

WAFER-SCALE LASER PANTOGRAPHY: II. LASER-INDUCED PYROLYTIC CREATION OF MOS STRUCTURES

BRUCE M. McWILLIAMS, IRVING P. HERMAN, RODERICK A. HYDE,
FRED MITLITSKY and LOWELL L. WOOD

Physics Department, L-278
Lawrence Livermore National Laboratory,
Livermore, CA 94550
(415) 422-1132

ABSTRACT

Laser-driven pyrolytic deposition, doping and etching of micron-scale structures has been studied using an argon-ion laser. By varying the appropriate gas mixtures over a p-silicon <100> substrate coated with a a thin film of silicon dioxide, the deposition of metal and polysilicon, the doping of silicon, and the etching of silicon dioxide and silicon were all successfully performed to create working metal-oxide-semiconductor (MOS) devices.

---

*Work performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under contract number W-7405-ENG-48.

WAFER-SCALE LASER PANTOGRAPHY: II. LASER-INDUCED PYROLYTIC CREATION OF MOS STRUCTURES

BRUCE M. McWILLIAMS, IRVING P. HERMAN, RODERICK A. HYDE,
FRED MITLITSKY and LOWELL L. WOOD.

Physics Department, L-278
Lawrence Livermore National Laboratory,
Livermore, CA 94550
(415) 422-1132

SUMMARY

This paper reports on a series of processes for generating micron-scale metal-oxide-semiconductor (MOS) devices suitable for fabricating ultra-large scale integrated (ULSI) circuits. The approach involves using a laser to locally heat a portion of the p-silicon <100> substrate coated with a thin film of silicon dioxide. Depending on the gas present above the substrate, this laser-induced pyrolysis results in the deposition of metal or polysilicon, doping of silicon, or etching of the substrate.

A theory is presented which gives quantitative predictions for each of the processes examined experimentally, for arbitrary laser beam pulse width and focal diameter. The temperature profiles inside the surface are found for each process, starting from the known beam profile. A Green's function technique is used when there is no phase change. When melting occurs, a 2-D time-dependent heat transfer code is used instead. In all of the processes, the gas phase diffusion of reactants to the surface, as well as the surface interactions, are modelled. To predict the doping profiles, diffusion within the substrate is also treated.

By sweeping or otherwise time-modulating a micron-scale laser beam spot on a time scale of the order of 10 ns the theory predicts that the heat transfer is sufficient to confine deposition, etching, or doping to within a micron spot diameter on a silicon substrate.

This localization should be with well-defined bounds due to the exponential dependence of the reaction rates with surface temperature over the range of interest. Substrates of lower thermal diffusivity (*e.g.*, silicon dioxide) are tolerant of even more prolonged exposure before 'thermal transport blurring' of the laser beam spot ensues.

In the described experiments, the output beam of a cw argon-ion laser (5145 Å) is amplitude-modulated by an electro-optic switch and then focused onto the silicon wafer to a spot size of $<$1-4 μm by an objective which also forms part of a wafer-viewing microscope. The wafer is situated in a reaction cell in an atmosphere of the appropriate neat gas or gas mixture (e.g., WF_6, HCl, SiH_4, PH_3, etc.).

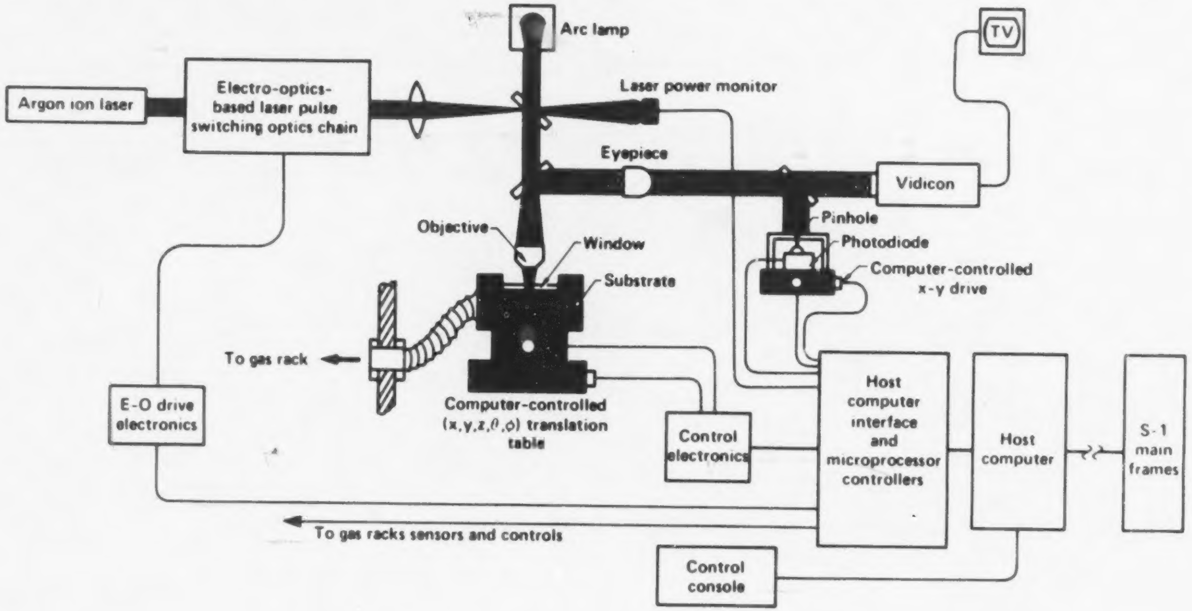
The target wafer is mounted in a stainless steel chamber, with a sapphire entrance window, that is directly plumbed to the gas handling/vacuum system. This reaction chamber is mounted on 0.5μm resolution x-y stepper motor translation stages, to control spatial positioning. An LSI-11/microcomputer interface allows dynamic control of the incident laser pulse properties (intensity, pulse, width, frequency, number of pulses per step, number of over-writes) with motion of the x-y stage, to permit the programming of direct laser write processes with a wide variety of possible duty cycles.

Results presented include working n-MOS transistors created by the laser-induced direct writing process described. Threshold voltages were shown to scale linearly with silicon dioxide thickness as expected. The process has been used to make NAND, NOR, and NOT gates, by wiring together multiple transistors in various fashions. Polysilicon lines with widths as narrow as 0.8 microns were deposited with a laser spot size of 1.8-2.0 microns, which demonstrated theoretical assertions that the minimum feature sizes possible should be smaller than the beam diameter. This implies that with state-of-art optics and positioning equipment, creation of circuits with 0.5 micron feature sizes will be possible using this technology with visible laser, while 0.2-0.3 micron feature sizes should be obtained with UV lasers.

Manufacture of Integrated Circuits by Laser-Induced Micro-Reactions

- Complete patterning of IC directly without resist or masks.
- Single environment from start to finish.
- Comprehensively computer-controlled process.
- Possibility of making "on-line" changes.
- Overnight turn-around times on new IC designs.

LASER MICRO DEPOSITION EXPERIMENT


Arc lamp
TV
Argon ion laser
Electro-optics-based laser pulse switching optics chain
Laser power monitor
Eyepiece
Vidicon
Objective
Window
Pinhole
Photodiode
Computer-controlled x-y drive
Substrate
To gas rack
Computer-controlled (x,y,z,θ,φ) translation table
E-O drive electronics
Control electronics
Host computer interface and microprocessor controllers
Host computer
S-1 main frames
To gas racks sensors and controls
Control console

Laser-Induced Pyrolytic Deposition Processes

- A focused laser beam on substrate induces chemical reactions of the form:

$$[MX_n]_{gas} \rightarrow [M]_{solid} + n[X]_{gas}$$

- Reaction rates are generally given by:

$$R \sim A \exp[-T_A/T]P^n$$

where T = substrate temperature, P is the reactant gas pressure, and T_A is the activation temperature.

- Dimension of deposit (ρ_0) relative to laser spot size (r_0) can be shown to be roughly given by:

$$\frac{\rho_0}{r_0} \simeq \sqrt{\frac{T}{T_A}}\sqrt{\frac{2}{(1-\frac{99^\circ K}{T_0})\ln(1+\frac{T-T_0}{T_0-99^\circ K})}} \sim \sqrt{\frac{T}{T_A}} \sim 0.3-0.4$$

for reactions of interest

- Experiments confirm that diameter of deposit can be made smaller than laser spot size by at least a factor of 2.

Typical Laser-Driven Material Deposition Reactions

- Tungsten deposited by:

  $WF_6 + 3H_2 \rightarrow W + 6\ HF$

- Nickel deposited by:

  $Ni(CO)_4 \rightarrow Ni + 4CO$

- Silicon deposited by:

$$\begin{aligned} SiH_4 &\rightarrow Si + 2H_2 \\ Si_2H_6 &\rightarrow 2Si + 3H_2 \\ Si_3H_8 &\rightarrow 3Si + 4H_2 \end{aligned}$$

- Doped polysilicon lines with conductivities as low as 10^{-2} to 10^{-3} ohm cm obtained by adding PH_3 to silane mixtures.

Laser-Driven Erasing of MOS Structures

- Etching of Silicon

$$\mathrm{Si} + 4\mathrm{H\,Cl} \rightarrow \mathrm{Si\,Cl_4}\uparrow + \mathrm{H_2}\uparrow$$
$$\mathrm{Si} + 2\mathrm{Cl_2} \rightarrow \mathrm{Si\,Cl_4}\uparrow$$

- Removing Oxide

$$\mathrm{SiO_2} + \mathrm{Si}\,|_{\mathrm{molten}} \rightarrow 2\mathrm{SiO}\uparrow$$

Fabrication of MOS Transistors and Small-Scale Integrated Circuits

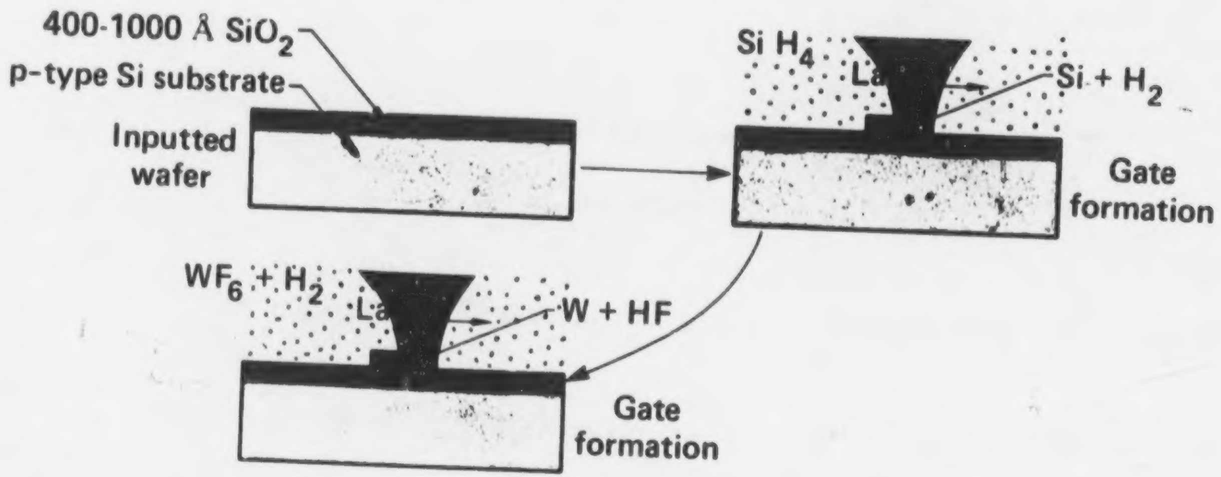
- Functional MOS transistors fabricated by direct-write sequence of laser processes.

- With 1000A gate oxide, threshold for transistors is observed to be 3 to 5 volts.

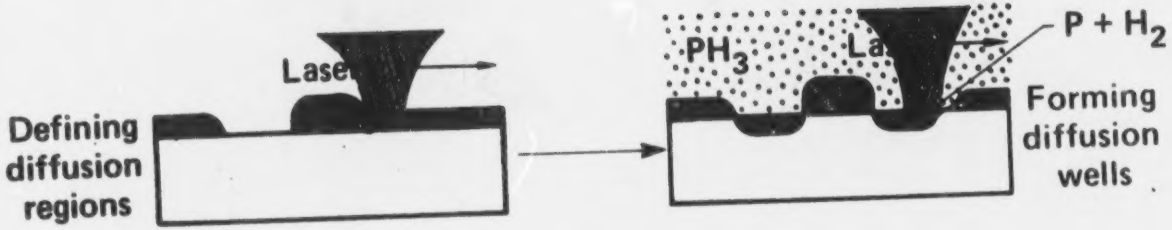
- Early devices had feature sizes of 2 to 4 microns.

- NOT, NAND and NOR gates fabricated.

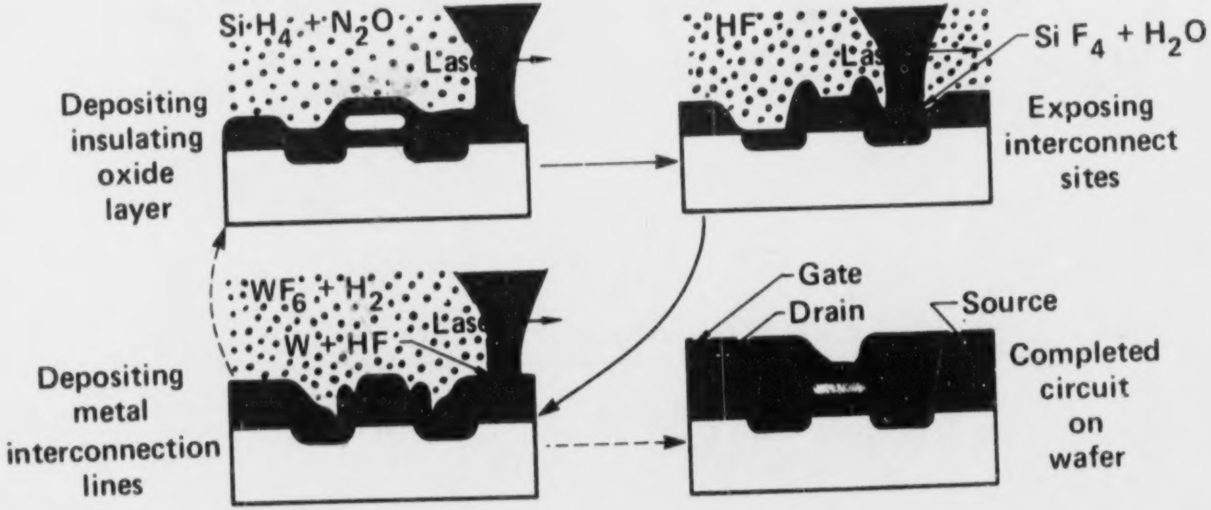
MOS TRANSISTOR GATE FORMATION BY LASER PANTOGRAPHY


400-1000 Å SiO_2
p-type Si substrate
Inputted wafer
$Si H_4$
La
$Si + H_2$
Gate formation
$WF_6 + H_2$
La
W + HF
Gate formation

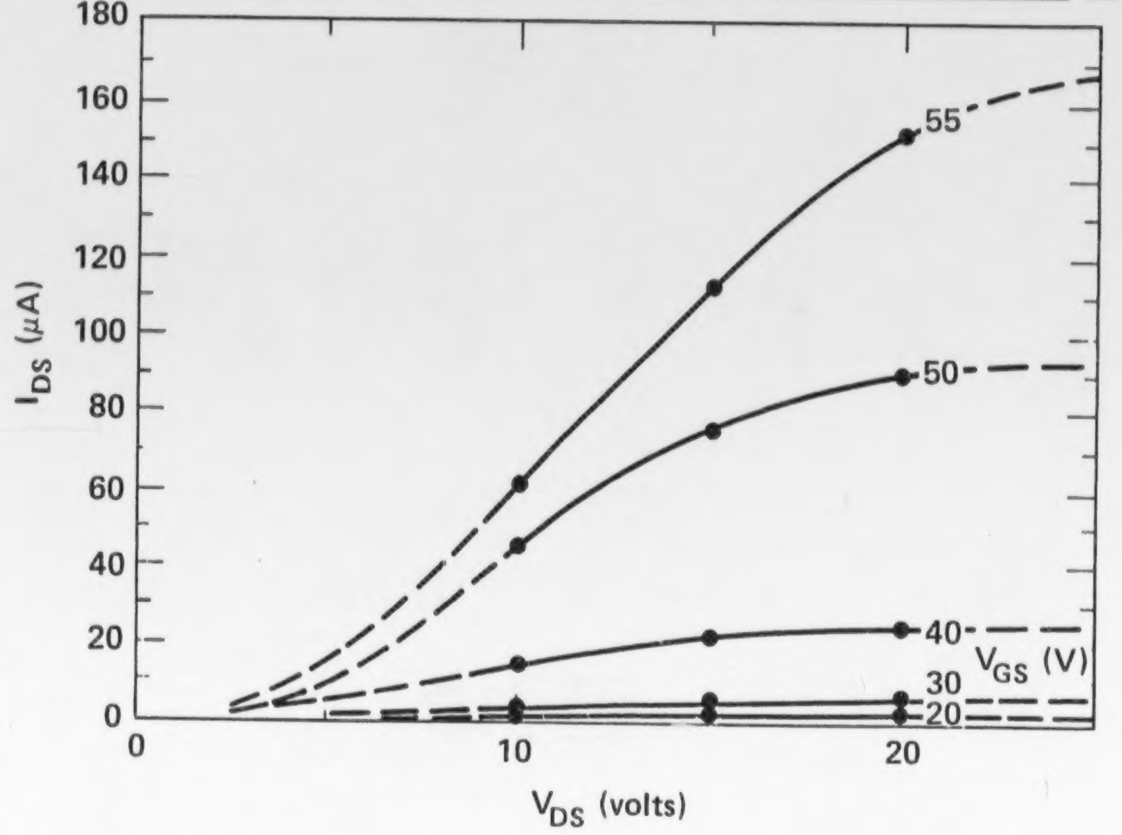
MOS TRANSISTOR SOURCE AND DRAIN FABRICATION BY LASER PANTOGRAPHY

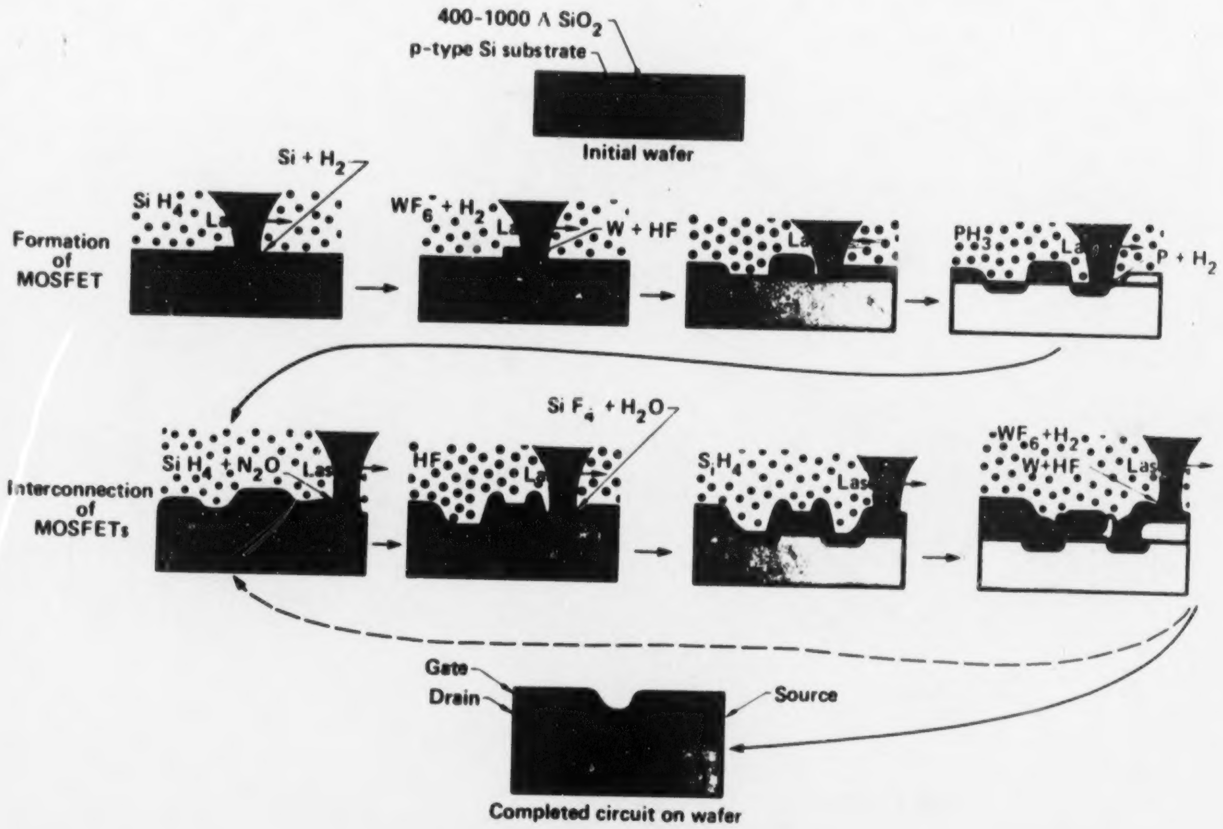

Defining
diffusion
regions
Lase
PH3
La
P + H2
Forming
diffusion
wells

FINAL STEPS IN MOS TRANSISTOR FABRICATION BY LASER PANTOGRAPHY



$Si H_4 + N_2O$
Laser
Depositing insulating oxide layer
HF
Laser
$Si F_4 + H_2O$
Exposing interconnect sites
$WF_6 + H_2$
Laser
$W + HF$
Depositing metal interconnection lines
Gate
Drain
Source
Completed circuit on wafer


180
160
140
120
100
80
60
40
20
0
I_{DS} (μA)
0
10
20
V_{DS} (volts)
55
50
40
30
20
V_{GS} (V)


400-1000 Λ SiO_2
p-type Si substrate
Initial wafer
Formation of MOSFET
SiH_4
$Si + H_2$
$WF_6 + H_2$
$W + HF$
PH_3
$P + H_2$
Interconnection of MOSFETs
$SiH_4 + N_2O$
HF
$SiF_4 + H_2O$
SiH_4
$WF_6 + H_2$
$W + HF$
Gate
Drain
Source
Completed circuit on wafer

MOS Transistor creation by laser pantography

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government thereof, and shall not be used for advertising or product endorsement purposes.

WAFER-SCALE LASER PANTOGRAPHY: III. FABRICATION OF n-MOS TRANSISTORS AND SSI CIRCUITS BY DIRECT-WRITE LASER-INDUCED PYROLYTIC REACTIONS

Bruce M. McWilliams, Irving P. Herman,
Fred Mitlitsky, Roderick A. Hyde, and
Lowell L. Wood

Department of Physics
University of California
Lawrence Livermore National Laboratory
Livermore, CA 94550

ABSTRACT

A complete set of processes sufficient for manufacture of n-MOS transistors by a laser-induced direct-write process has been demonstrated separately, and integrated to yield functional transistors. Gates and interconnects were fabricated of various combinations of n-doped and intrinsic polysilicon, tungsten, and tungsten silicide compounds. Both 0.1 μm and 1 μm thick gate oxides were micro-machined with and without etchant gas, and the exposed p-Si [100] substrate was cleaned and, at times, etched. Diffusion regions were doped by laser-induced pyrolytic decomposition of phosphine followed by laser annealing. Along with the successful manufacture of working n-MOS transistors and a complete set of elementary digital logic gates, this letter reports the successful use of several laser-induced surface reactions that have not been reported previously.

---

*Work performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under contract number W-7405-ENG-48.

Direct laser-writing and processing of semiconductor surfaces for electronic applications is becoming recognized as a versatile and powerful technology.[1] Lasers have been utilized to etch,[2,3] dope,[4,5] and deposit[6-9] materials on crystalline silicon substrates interfacing gaseous reactants. Both photolytic (gas-phase photodissociation)[8] and pyrolytic (surface heating)[5] techniques have been used to demonstrate the sub-micrometer spatial resolution obtainable with direct laser-writing. This letter reports the first fabrication of transistors using solely direct laser-write techniques ("laser pantography"). These n-MOS transistors and small-scale integrated circuits were made using several new, as well as some previously-demonstrated, direct laser-writing, pyrolytic methods. The fabrication cycle and preliminary results on early transistor and digital logic devices are reported here.

The experimental apparatus used in these studies is diagrammed in Fig. 1. A CW argon-ion laser (5145 Å) is amplitude-modulated (with typically 20 ns - 10 ms pulses) by an electro-optic switch positioned between cross-polarized Glan-Thompson prisms. The temporally-controlled laser beam is then imaged on the semiconductor substrate by a 6 cm focal length lens followed by an 28 X, 0.45 N.A. microscope objective. At times, arc lamp light also irradiates the substrate through this objective and projects an image onto a vidicon through an eyepiece, in this expanded microscope arrangement, to support direct viewing of the surface or impinging laser beam. This magnified laser beam spot is also projected on an x-y scanning pinhole-photodiode assembly to permit spot size measurement of the focused beam. The typical focused beam spot diameter (1/e intensity points) in these experiments was 1.9 μm.

The target wafer is mounted in a stainless steel chamber, topped with a sapphire entrance window, which is directly plumbed into the gas handling/vacuum system. This reaction chamber is mounted on a pair of 0.5 μm resolution x-y stepper motor-driven translation stages, which control the spatial position of the wafer. An LSI-11/23 microcomputer interface provided dynamic control of the incident laser pulse properties (intensity, pulse width, frequency, number of pulses per processing site, number of processing cycles) and also controlled motion of the x-y translation stage, to permit the programming of direct laser write processes over a wide variety of possible process parameters.

The target substrates were p-type [100], 3.4 - 4.6 Ω-cm crystalline silicon with a predeposited, uniform silicon dioxide film on the surface. To date, substrates with either 1 μm oxide grown by a wet thermal oxidation process or 900Å oxide grown by a dry thermal oxidation process have been used to fabricate MOSFETs. Before installation in the reaction chamber, the wafer is sequentially immersed in boiling solutions of trichloroethane, acetone, methanol, and de-ionized water in a clean room environment. The wafer is then mounted in the reaction chamber, and baked out, under high vacuum. The final surface preparation step is *in situ* ozone etching; the ozone is produced by flowing oxygen through the reaction chamber while irradiating with UV radiation from a Hg arc lamp. This is followed by high vacuum bake out of the entire cell. Prior to introducing each new reagent gas for use in the reaction cell, the vessel is flushed several times with helium or nitrogen, and is then passivated with a sample of the new reactant. All reagents were used as supplied by the manufacturer without further purification, except for pumping off possible volatiles at 77°K as required.

The process developed for MOS transistor fabrication is diagrammed in Fig. 2. A conducting gate structure of dimension typically 1-5 μm wide and 10-30 μm long is first deposited on the oxide and "wired" to a probing pad. These pads are typically 30-40 μm squares that are made by the same process as the gate. Diffusion regions are patterned by removing the oxide in two rectangular regions ($\sim$30 × 40 μm) abutting either side of the gate. This is sometimes followed by etching the exposed silicon. Removing the oxide to make substrate electrical ground is also accomplished in this step. The source and drain are appropriately n-type doped with phosphorous, pulse-laser annealed, and then attached to probing pads. Device characteristics are currently measured after the wafer has been removed from the chamber.

Several alternate technologies have been tested for each step outlined above. These are briefly described here, and will be further detailed elsewhere.

For the gate structures and interconnects, doped and undoped polycrystalline silicon, tungsten, and tungsten silicides were examined. The most successful structures were formed by depositing polysilicon using a silane/disilane/phosphine mixture; this was sometimes followed by forming a tungsten overlayer. Typically, rapid deposition of highly localized features was accomplished by first irradiating with relatively high laser powers (0.5-1.0 W) to seed the surface, and then over-writing the deposits several times at lower laser powers to obtain the desired film thickness.

Silicon deposition was studied using various mixtures of silane,[6,7] disilane, and trisilane at 100-800 torr, sometimes laced with phosphine (0.01-0.5%). The temperature thresholds for deposition decreased with increasing silane chain length, as predicted by the CVD experiments of Gau, et al.[10] Disilane surface reactions produced deposits at laser powers as low as $\sim$0.1 W (focused to 1.9 μm beam diameter). The rate of deposition was on the order of 0.5-5 mm/s depending on laser power and gas pressure. Using trisilane, silicon features were deposited much faster than with disilane; however, they were much less localized.

Localization of the silicon deposits decreased with increasing laser power and increasing integrated irradiation times. The approximate deposition rate using silane was a factor of 2-5 below that of disilane (at similar pressures), but silane-deposited silicon was usually more localized. Polysilicon linewidths as narrow as 0.9 μm were deposited onto 0.1 μm thick oxide by the silane reaction. Such enhanced localization (linewidths ~2.5 times smaller than the effective laser beam diameter) is predicted by theory.[11]

Doped silicon was deposited by adding small levels of phosphine (0.01-0.5% PH_3) to the 800 Torr silane. The relatively rapid surface deposition of phosphorous on the oxide apparently increased the nucleation barrier over that of undoped silicon, consistent with the observation of Everstyn and Put.[12] Lines with resistivities of $\sim 10^{-2}$ - 10^{-3} Ω-cm were deposited in this fashion.

Tungsten lines were deposited by the laser-induced pyrolytic hydrogen reduction of tungsten hexafluoride.[9] Experiments were performed with a 3:1 - 4:1 ratio of hydrogen : tungsten hexafluoride, with the pressure totalling 800 Torr. This reaction proceeded at laser beam powers as low as ~0.1 W. When tungsten deposition was initiated with higher laser beam powers (0.6-0.8 W), using 1-10 μsec pulses, the deposited material appeared shiny and metallic.

Silicon dioxide etching by local laser heating was examined in vacuum, in H_2 and in H_2-buffered HCl. Sufficiently rapid and localized oxide removal occurred by locally heating the oxide (by silicon substrate absorption followed by heat transfer to the surface) in vacuum, with an etch rate of ~40 μm/sec for ~3.5 W incident laser power. No damage to the surrounding oxide was observed by optical means. The oxide removal mechanism may involve the known decomposition of SiO_2 on Si at temperatures above 900°C to form the volatile SiO.[13] With use of 180 Torr HCl/540 torr H_2 over the substrate, the etch rate was increased to ~100 μm/sec with only 0.6 W incident beam power. After oxide removal, the exposed silicon substrate was either cleaned by laser-heating (desorption) in vacuum or etched by pyrolytic reactions with H_2-buffered HCl.[2]

The source and drain regions of MOSFETs were doped by first either depositing a phosphorous surface layer (0.5 Torr PH_3/800 Torr He) or by laying down a thin layer of doped polysilicon (PH_3/SiH_4). Afterwards, the dopants were driven in by annealing with 1 μs-10 μs duration pulses of 1-2 W power. These intensity levels were sufficient to cause transient local melting, and changes in the color of the source and drain regions were observed which were attributed to recrystallization of previously amorphous silicon.

MOSFET devices were made with both 1 μm and 0.09 μm thick gate oxides. The DC

I-V properties of a 1 μm thick gate oxide transistor are shown in Fig. 3. The ~30 V gate threshold is that predicted by theory for this thick gate oxide field-effect transistor. Gate thresholds for MOSFETs with 0.09 μm gate oxide thickness were in the 3-5 volt range. The leakage current levels in these early devices significantly affected device performance at higher source-drain voltages.

Small sets of these MOSFETs have been interconnected to realize NOT, NOR and NAND gates, the latter with two inputs each. One of these is shown in Figure 4.

Studies are continuing to better characterize and control the different process steps for ultra-large scale integrated ULSI circuit prototyping by the mask-free, resist-free, direct laser writing technique described here. This technology allows for complete fabrication, testing, and repair of ULSI circuits in a single environment. The system of Figure 2 is presently being upgraded to support extremely rapid (~10^3/sec) fabrication interconnection and testing of n-MOS transistor-based ULSI circuits with sub-micron minimum feature sizes.

Acknowledgment

The authors wish to thank H. W. Chin and A. H. Weisberg for valuable discussions, and L. R. Mendonca, A. E. Ormsby, and D. E. Duerre for expert technical assistance.

REFERENCES

1. R. M. Osgood Jr., "Laser microchemistry and its application to electron-device fabrication," Ann. Rev. Phys. Chem. **34**, (1983) (in press).
2. D. J. Ehrlich, R. M. Osgood Jr., and T. F. Deutsch, Appl. Phys. Lett. **38**, 1018(1981).
3. F. A. Houle and T. J. Chuang, J. Vac. Sci. Technol. **20**, 790(1982).
4. T. F. Deutsch, D. J. Ehrlich, D. D. Rathman, D. J. Silversmith, and R. M. Osgood Jr., Appl. Phys. Lett. **39**, 825(1981).
5. D. J. Ehrlich, and J. Y. Tsao, Appl. Phys. Lett. **41**, 297(1982).
6. D. J. Ehrlich, R. M. Osgood Jr., and T. F. Deutsch, Appl. Phys. Phys. Lett. **39**, 957(1981).
7. D. Bauerle, P. Irsigler, G. Leyendecker, H. Noll, and D. Wagner, Appl. Phys. Lett. **40**, 819(1982).
8. D. J. Ehrlich, R. M. Osgood Jr. and T. F. Deutsch, J. Vac. Sci. Technol. **21**, 23(1982).
9. S. D. Allen and A. B. Tringubo, J. Appl. Phys. **54**, 1641(1983).
10. S. C. Gau, B. R. Weinberger, M., Akhtar, Z. Kiss, and A. G. MacDiarmid, Appl. Phys. Lett. **39**, 436(1981).
11. I. P. Herman, R. A. Hyde, B. M. McWilliams, A. H. Weisberg, and L. L. Wood, Materials Research Society Symposia Proceedings (November 1982), **17**, *Laser Diagnostics and Photochemical Processing for Semiconductor Devices*, R. M. Osgood Jr., S. R. J. Brueck, and H. R. Schlossberg, (editors), Elsevier Science Publishing Co., 9(1983).
12. F. C. Eversteyn, and B. H. Put, J. Electrochem. Soc. **120**, 106(1973).
13. J. C. Bean, in *Impurity Doping Processes in Silicon*, F. F. Y. Wang editor, p. 189, North-Holland, New York, 1981.

Figure Captions

Figure 1: The experimental set-up for direct-laser writing, see text for further details.

Figure 2: Schematic of the steps involved in MOS transistor fabrication by laser pantography, as described in the text. After cleaning the inputted wafer (p-type Si with SiO_2 overlayer), gate structures are formed by depositing polysilicon and tungsten. The diffusion regions are defined by removing the oxide; the locally bare silicon substrate is cleaned and sometimes etched (not shown), and appropriately doped with phosphorous. Metal interconnection lines (and probing pads, as needed) are then deposited. To ensure electrical isolation, in a refined version of this last step currently under development, there are instead several repeated sequences of thick (>1 μm) field oxide deposition, followed by oxide removal at the desired interconnect sites and deposition of metal interconnects.

Figure 3: The DC I-V characteristics of a laser-fabricated n-MOSFET with 1 μm thick gate oxide, with drain-source current plotted versus drain-source voltage for a sequence of gate-source biasing voltages.

Figure 4: A two-input NOR gate, fabricated with 2 parallel MOSFETs whose gates are the A and B inputs, whose sources are tied to ground and whose drains are tied together to a pull-up FET at the NOR output point.

LASER MICRO DEPOSITION EXPERIMENT

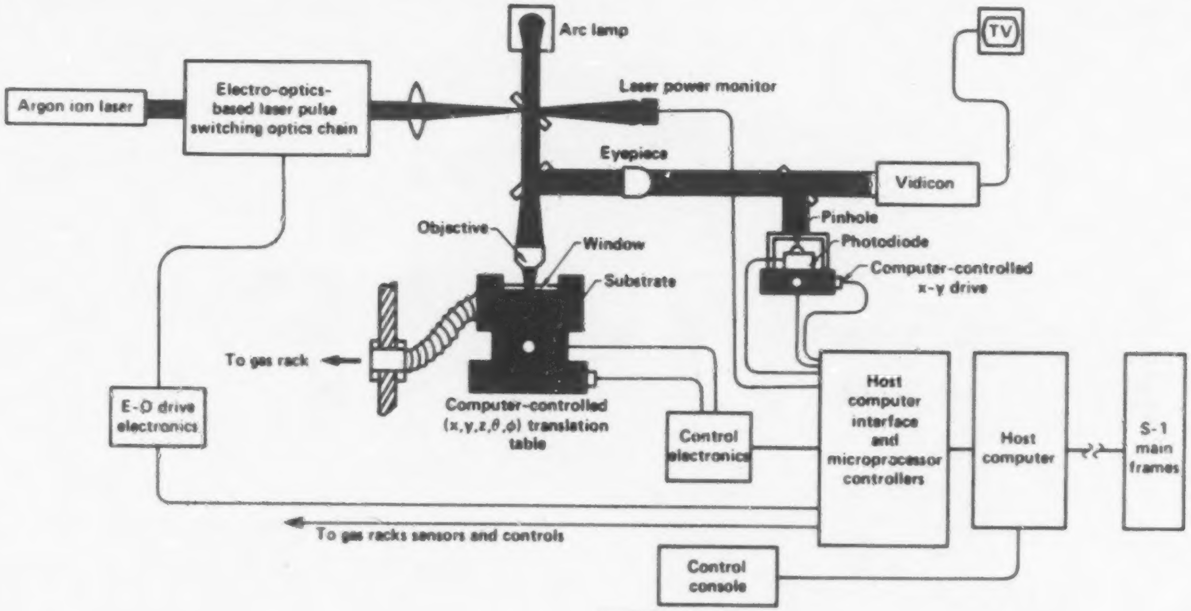
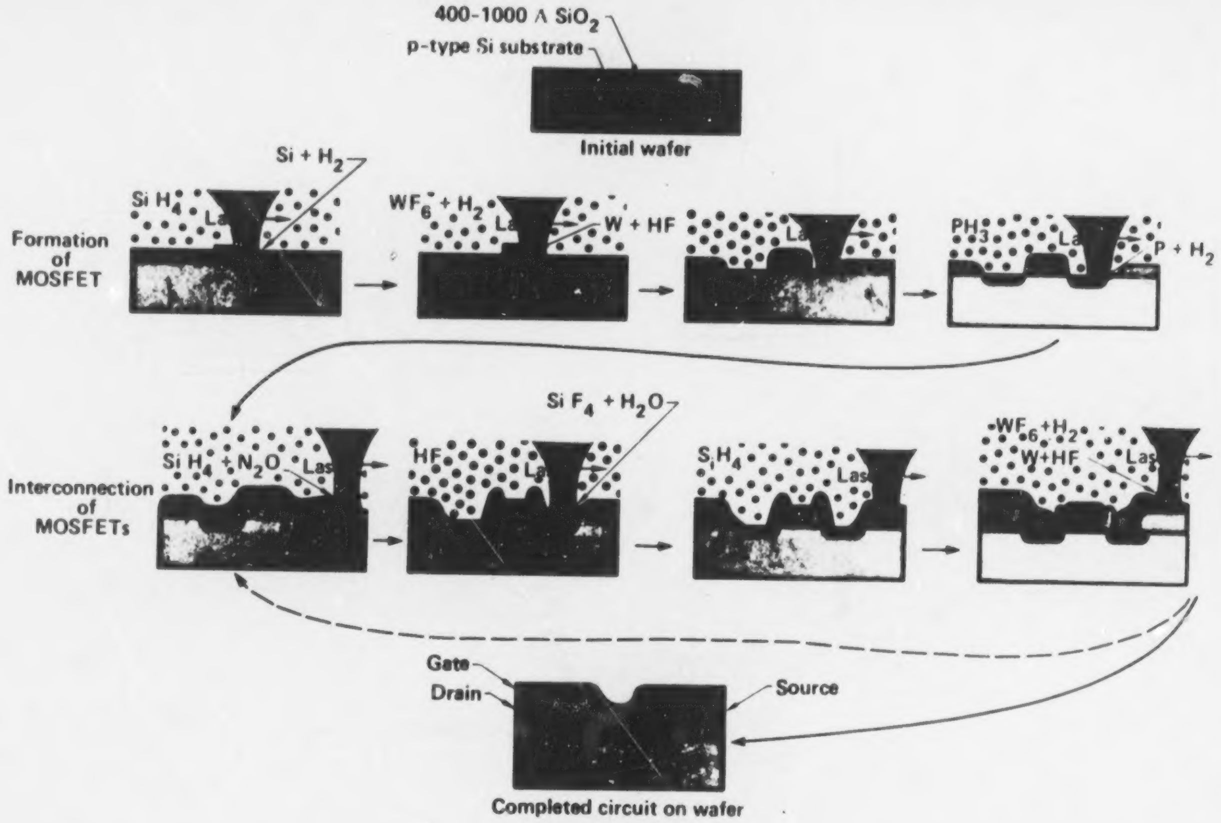
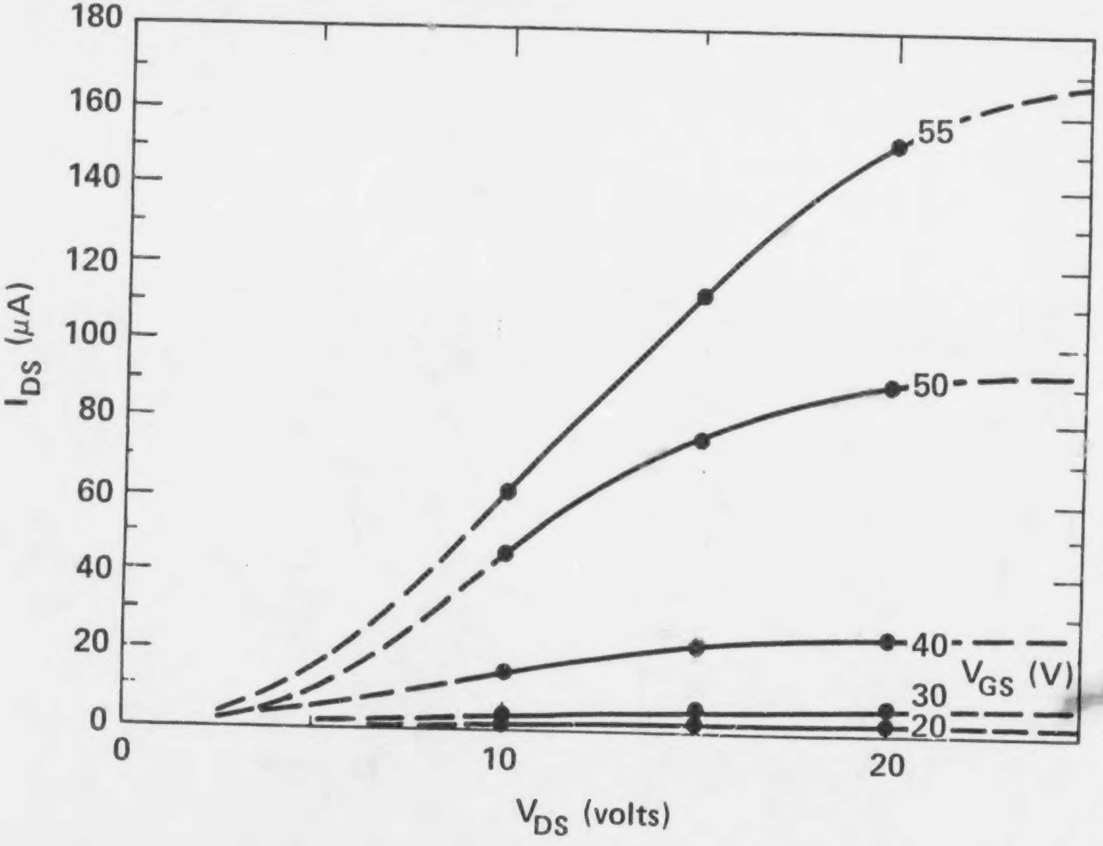

Arc lamp
Argon ion laser
Electro-optics-based laser pulse switching optics chain
Laser power monitor
Eyepiece
Vidicon
TV
Objective
Window
Substrate
Pinhole
Photodiode
Computer-controlled x-y drive
To gas rack
Computer-controlled (x,y,z,θ,φ) translation table
E-O drive electronics
Control electronics
Host computer interface and microprocessor controllers
Host computer
S-1 main frames
To gas racks sensors and controls
Control console

Figure 1


400-1000 Λ SiO_2
p-type Si substrate
Initial wafer
Formation of MOSFET
Si H_4
La
Si + H_2
WF_6 + H_2
La
W + HF
La
PH_3
La
P + H_2
Interconnection of MOSFETs
Si H_4 + N_2O
Las
HF
La
Si F_4 + H_2O
S_iH_4
Las
WF_6 + H_2
W+HF
Las
Gate
Drain
Source
Completed circuit on wafer

MOS Transistor creation by laser pantography

Figure 2
Expanded



180
160
140
120
100
80
60
40
20
0
I_{DS} (μA)
0
10
20
V_{DS} (volts)
55
50
40
V_{GS} (V)
30
20

Figure 3

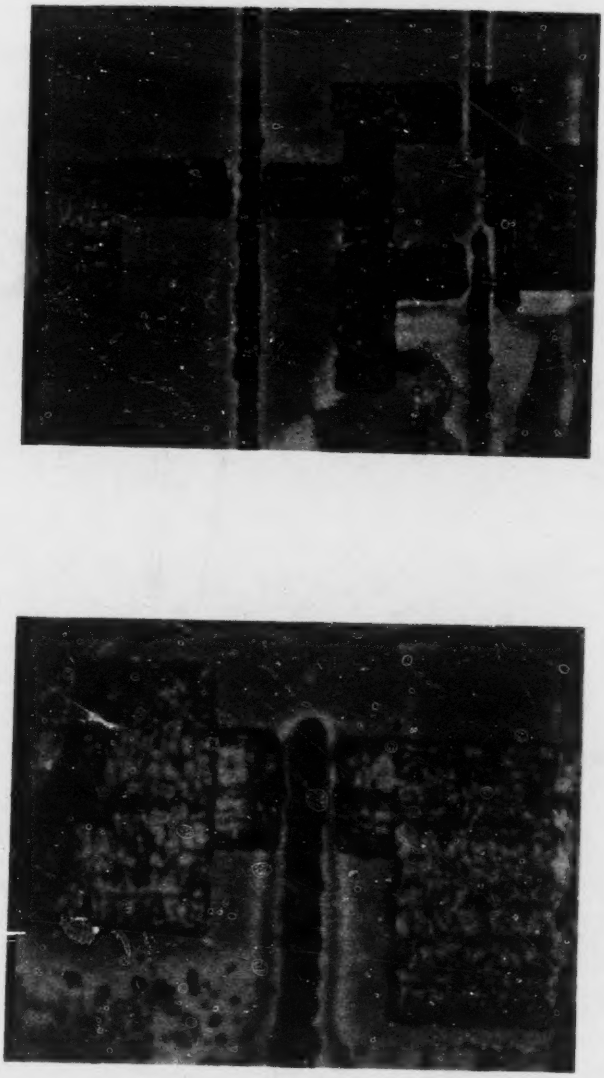


Figure 4

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government thereof, and shall not be used for advertising or product endorsement purposes.

LLL-PROP 185

SUPERCOMPUTING RESEARCH AND DEVELOPMENT: THE LLNL PROGRAM AND PROPOSALS

C. B. Tarter
M. May
C. Haussmann

March 1983

Lawrence
Livermore
Laboratory

EXECUTIVE OVERVIEW

After their advent in the late 1940s, the performance of the leading-edge digital computers more than doubled every other year, on the average, through the late 1960s. In the decade and half since that time, only a threefold performance gain in scientific computing capability has been realized, though another several-fold may become commercially available in the latter part of the 1980s. By any measure, the pace of advance in this area has slowed drastically. This has resulted more from a lack of vigorous application of technical advances than from a deficiency in the technology support base.

Historically, the U.S. has maintained its technical leadership by rapid advances in microelectronics that were commercially attractive to convert to gains in supercomputing power. This "business as usual" approach is no longer effective. The supercomputer of the future will have to be developed by manufacturers, computer scientists, and applications people working together, with the government providing incentives. Design advances in one field will need to be implemented in another. Attention to potential applications will be required to exploit the power of the new machines.

Only a coordinated and focused national program can accomplish these objectives. This conclusion has clearly been reached by Japan, which has launched an integrated national program to advance the performance of scientific computer systems by a factor of 1000 by the end of the decade. The Japanese have also initiated a major effort to exploit the full application potential of those systems, their so-called fifth generation project.

The implications for the economy and defense posture of the United States of a loss of leadership in the field of major computing systems would be profound. The world-wide information processing industry last year involved a cash stream aggregating over $100 billion, most of which flowed through U.S. hands. Loss of a significant fraction of it would adversely affect basic American economic interests, a direct impact. Indirectly, the applications of orders of magnitude more capable computers, available at reduced total costs, could provide an economic edge to many industries. With regard to defense issues, a Defense Science Board Study has identified many technological areas likely to have revolutionary impact on military posture in the 1990s which are predicated on advanced computing/microelectronics capabilities. Loss of the U.S. technological edge in computing to any foreign country could have grave military consequences.

A definitive U.S. initiative to maintain our computing supremacy therefore seems imperative. This initiative must include regaining across-the-board technological supremacy in all aspects of the development and production of advanced components and integrated computing systems, and the continuing endowment of such systems with software constructs which advance as rapidly as possible the efficiency with which people use these systems. More vigorous and more focused pursuit of applications of these advances is also required. To accomplish these goals, an interrelated effort is needed among industry, government, and universities. Computer and microelectronics firms must supply the hardware skills, universities the fundamental research and new talent, and the federal agencies and laboratories the integration and applications of the other thrusts.

The Lawrence Livermore National Laboratory (LLNL) has pioneered in all aspects of computer systems integration, utilization and creation, and continues to be one of the most computing-intensive technical centers in the world. LLNL's current efforts are driven by, and focused on our own programmatic needs. We are prepared to expand them in pertinent areas, and to take a leadership role in an American national supercomputing program.

In this paper we describe LLNL's activities in supercomputing research and development, emphasizing areas most relevant to a broad national program. We give our recommendations for the goals and scope of such a program, and then discuss our own needs and major efforts. These include an assessment of future programmatic requirements in computing; a description of LLNL's computational environment; a synopsis of software development activities; the status of the S-1 supercomputer project; and related work in microelectronics. We then point out potential applications to scientific and defense programs outside our direct interest. Finally we discuss the incremental budgeting requirements associated with our proposals. These include $10-15M in annual operating funds plus $10-15M in initial capital expenditures. We believe these estimates to be realistic for our participation in a vigorous, but not "all-out", national program whose overall level is about an order of magnitude larger than LLNL's proposed effort.

SUPERCOMPUTING RESEARCH AND DEVELOPMENT:

THE LLNL PROGRAM AND PROPOSALS

Table of Contents

	Page
Executive Overview	(i)
I. Introduction	1
II. Goals and Scope of a National Supercomputing Program	1
III. Supercomputing Research and Development at LLNL	3
A. Programmatic Needs	4
B. Computational Environment	5
C. Software Development	7
D. The S-1 Project	9
E. Component Technologies	11
IV. Applications to Other Scientific and Defense Programs	12
A. Interactions with the National Computer Community	13
B. Defense and Systems Applications	14
V. Budget and Resource Matters	15
VI. Appendices	17
A. Computational Needs of the Nuclear Weapons Program	17
B. Computers at LLNL: History, Status, and Plans	17
C. Software Research for Supercomputing	20
D. The S-1 Project	29
E. Expanded Microelectronics Facilities	36
F. Laser Pantography	38
G. Monolithic Implementation of Supercomputers	40
VII. Bibliography	50

I. INTRODUCTION

The purpose of this paper is to summarize LLNL's views on the goals and scope of a national supercomputing effort, and to outline our programs and proposals in supercomputer research, development, and use. We describe our current and recommended work in relevant technical areas, ranging from conventional acquisition and use of supercomputers to innovative developments in software and hardware. In most cases these activities are highly synergistic with those elements necessary for a national program aimed at dramatic improvements in scientific and symbolic computing power. They also have potential applications in the defense and general computer communities, and these possibilities are discussed. A series of Appendices provides more detail on each of the topics covered in the main text.

II. GOALS AND SCOPE OF A NATIONAL SUPERCOMPUTING PROGRAM

As emphasized by recent technical and defense panels, supercomputing research and development in the U.S. is at a critical point. These studies have also described the great impact of future supercomputers and their associated technologies on both national security and economic development.

Accepting the need for supercomputers, the question arises as to whether the historical process for producing such machines will function effectively in the 1980's. We believe the answer to be no. The first reason is economic: small and intermediate-scale computers are much more profitable product lines for manufacturers. The second difficulty is that uncoupled development of the required technologies will not suffice to expeditiously produce a sequence of ever more powerful digital computers. Almost all "models" of future supercomputers envision them as highly parallel machines whose effective use will require as much innovation in software as in hardware. Consequently, the path to a more powerful supercomputer is not guaranteed by any inexorable technological logic.

Only a coordinated national program offers a realistic expectation of attaining a dramatic and rapid increase in supercomputer capabilities. This conclusion has clearly been reached by Japan, which has initiated a program involving tight collaboration among government, industry, and the universities.

A consensus of the recent review groups is that an increase in supercomputing power of several orders of magnitude is technically feasible by the end of the century. The Japanese goal is a factor of 1000 over the next decade. Perhaps a better and more realistic goal for the United States would be a factor of 100 growth in each of the next two decades (The U.S. commercial goal appears to be about 10-fold growth in the next 6-8 years). To accomplish this objective will require significant progress in each of the following areas:

- Component technology--integrated circuits, computer-aided tools for circuit design, rapid prototyping turnaround
- Computer implementation technology--packaging, testability, fault tolerance
- Expanded use of CAD/CAM for large computer systems
- Computer architecture
- System software for parallel computing--operating systems, compilers, high level languages
- Algorithms

The first two fields, which form the foundation for the physical realization of supercomputers, have a large overlap with the integrated circuit industry and with programs independently undertaken by computer manufacturers. A national supercomputing effort could bring additional resources and advanced concepts to these areas, while interfacing productively with them. The third area, the use of CAD/CAM for large computers, is essential in a fast-paced program in which hardware and software developments are coupled. It will be necessary to implement design changes reliably and with much more rapid turnaround than is currently typical in prototype systems. The remaining three topics are not effectively addressed by any current program, and their content will largely be unique to a well-integrated supercomputing effort. We briefly discuss them.

First, the national program strategy must recognize that the architecture of future supercomputers is a relatively open question, although some form of parallel architecture seems almost inevitable. An aggressive program to explore innovative architectures must be undertaken in conjunction with the effort to develop higher speed integrated circuits and advanced hardware implementation techniques.

Second, a major effort will be required to develop system software which supports efficient use of the new parallel architectures. Operating systems must be written to support efficient scheduling of large numbers of parallel processes, and high-level languages, compilers, and debuggers for parallel machines need to be developed. Ideally, these problems should be addressed through implementation of experimental software on a variety of simulated and prototype computers.

Third, the nature of the algorithms must change to take advantage of the parallelism inherent in advanced supercomputers. For machines with large numbers of processors this may require major conceptual changes in our approach to numerical simulation.

Finally, to maintain rapid progress it is necessary that all participants in a national program be able to readily evaluate each other's work and exchange ideas. A useful way to help accomplish this would be to have all research sites connected to a common national computer network. This would provide very small users, particularly those at universities, with equal access to the prototype and production facilities of the larger centers.

In summary, we believe that a national effort in supercomputing will need to tie together research in a number of fields in a coordinated way. The program will emphasize highly integrated circuitry, CAD/CAM for computer design and implementation, parallelism in architecture, innovation in algorithms and systems software, and the means to communicate advances rapidly through a national computer network.

III. SUPERCOMPUTING RESEARCH AND DEVELOPMENT AT LIVERMORE

LLNL's work on nuclear weapons has always involved intensive use of the most powerful supercomputers available. This has impelled us to develop operating systems, networks, and other systems software for many of our computers. Our weapons design codes have then been adapted relatively straightforwardly for new computers, taking full advantage of the speed of the new machines. In Sections A and B below we briefly describe the history of this process, and project its likely future trends. As is evident from the discussion, we believe this "conventional" path is unlikely to meet fully the long term computing needs of our programs. Consequently, we have initiated several new projects in software for advanced computers, and these are described in Section C. The status of LLNL's S-1 supercomputer project, and the potential use of its mainframes for research and/or production purposes, is discussed in Section D. A final section outlines some of our related work on semiconductor technology.

A. Programmatic Needs

Numerical modeling with large scale digital computers has long been an essential tool in the design of nuclear explosives, in analyzing and predicting atmospheric phenomena, in operations research and intelligence applications, and in other government and commercial endeavors. LLNL has been at the forefront of several of these applications of supercomputers. Our design and understanding of nuclear weapons have been strongly influenced by the available computing power, and the exploitation of current advanced weapons concepts will require machines much faster than the Cray-1. Progress on inertial confinement fusion has been similarly tied to digital computers, and numerical simulation of magnetic fusion plasmas has become increasingly important.

Future trends indicate an even greater role for supercomputers. The basic reason is clear: computing is a relatively inexpensive, flexible, and repeatable operation, while the experimental alternatives are becoming more expensive (and this disparity is rapidly increasing). For example, the cost of a typical nuclear weapons experiment at the Nevada test site substantially exceeds the cost of a Cray-1 computer. Analogously, the next generation of proposed magnetic fusion devices may be so costly that elaborate simulation could be the mandated mechanism for choosing between competing schemes.

In Appendix A (a Secret Restricted Data Document) we have projected the computational requirements of the nuclear weapons program over the next decade. Based on historical trends and our analysis of the required complexity of the numerical simulation, we have estimated that a 100-1000 fold increase in computing power will be necessary to explore and understand the full potential of our current advanced weapons concepts. Such an increase is primarily due to the geometrical and physics richness of those new weapon designs and their effects.

Other national programs for which LLNL has significant responsibility will also derive major economic and technical benefits from greatly increased computing power. In both inertial confinement and magnetic fusion it will be possible to more realistically model experimental configurations, and to eliminate many of the simplifying assumptions now made. Our atmospheric scientists will be able to do 3-dimensional simulations of macroscopic phenomena in less than real time, and to incorporate chemistry and other reactive mechanisms in their meteorological models.

In summary, an increase in effective scientific computing power of 2 to 3 orders of magnitude is essential to advanced nuclear weapons development. It is also one of the most cost-effective methods for achieving rapid progress in our major energy and environmental programs. As discussed later (Section IV) it has correspondingly high potential in other areas important to national security, science and technology.

B. Computational Environment

Since its beginning LLNL has acquired the most powerful computers available in order to execute the large physics modeling codes used to design nuclear weapons. As a result, on many occasions over the past 30 years, we have purchased the first or early serial number of a new line of computing equipment (a detailed list is given in Appendix B).

It is often true that the first few units of a new computer line are devoid of usable software for the large scale numerical modeling community. Thus, in the case of many of the new machine types, we have found it necessary to develop and provide the total operating environment plus all the software needed to adapt the computer to the local computing center. This effort typically included, besides the application codes, the operating system, highlevel programming languages, and interfaces for input/output and data storage.

During the mid-1960s LLNL adopted the practice of timesharing its largest computers in a network configuration. The Livermore Computer Center (LCC) currently provides a time-shared network access of 4 Cray-1s and 3 CDC 7600s to 2500 users at LLNL. The network is a rich environment for data storage and input/output as well, since over 3 x 10^{12} bits of memory are on-line to the major computers, and we produce about 13 million page-equivalents of output each month.

In addition to the LCC, LLNL developed and operates the National Magnetic Fusion Energy Computer Center (NMFECC). This national network, currently centered on two Cray-1s and a CDC 7600, services over 40 sites around the U.S. that are involved with research in magnetic fusion. This network uses state-of-the-art satellite communications for data transfer to/from its major sites. NMFECC played an important role with LCC in developing the initial Cray-1 operating system, and its current version (CTSS, the Cray Time Sharing System) has been largely adopted by the Los Alamos National Laboratory and the Air Force Weapons Laboratory.

As the next generation of computers develops in the 1980s, LLNL will need to acquire several of the best of them to carry out its programs effectively. Since these computers are likely to use some form of multiple processor architecture, it is important that timely research and development work be done to evaluate the various potential configurations. Hands-on experience is a valuable way to study the characteristics of a new machine and measure its effective performance level, and we believe that a national supercomputing program should have as an essential part the acquisition and support of state-of-the-art (prototype) computer systems for use by researchers. Government-sponsored purchase of these machines could give hardware manufacturers an additional incentive to invest developmental dollars on supercomputers which currently have limited market appeal. Small companies and inventors showing real promise would thereby have an opportunity to offer initial hardware implementations of their designs.

LLNL is in an excellent position to offer the proper environment for hosting some of these prototype supercomputer architectures. We have both the large scale applications and the experienced staff who can collaborate with like-minded workers from academic organizations, national laboratories, and industry. As cited above, we also have a proven track record for the installation and effective utilization of large computers over the past 3 decades.

Our goal would be to host and benchmark several research and/or new production machines. The S-1 is an obvious possibility because of its propinquity and the ability to carry out closely coupled hardware-software development. In addition several major computer manufacturers have shown substantial interest in its architecture and potential. We have made pilot studies of the feasibility of hosting other machines, and these possibilities are discussed in Appendix B. The particular choices would depend on the resources provided, the operational readiness of the machines, and the near-term goals of the associated software program.

C. Software Development

There is a wide array of software issues which we believe should be investigated. Nearly all of our current activities are in embryonic stages at modest resource levels. Several-fold increases will be necessary to achieve progress commensurate with that anticipated in hardware areas, and to guide advances in those areas. Some of the major software topics are discussed below; a more complete description is given in Appendix C.

Performance Evaluation of Machines: In order to quantitatively understand the performance of parallel computers, extensive measurements must be made on real machines and those still in planning stages must be simulated. To support this effort, we have in the past year created a first-generation simulator which allows the performance of real application codes on Multiple Instruction/Multiple Data (MIMD) machines to be estimated. Additionally, we have performed measurements on the Cray X-MP and the Denelcor HEP. Our efforts need to be significantly expanded in:

- Simulator development and validation
- Systematic analysis of parallel processor performance
- Benchmarking -- local and remote

Parallel Algorithms: The highly parallel nature of future supercomputers will eventually require algorithms for numerical simulation to be redesigned using large numbers of cooperating and communicating parallel processes. Loss of efficiency due to frequent synchronization or interprocess communication must be minimized. Our approach to these problems is twofold. First, we have begun construction of a set of benchmark problems appropriate for our applications codes, and we are studying ways to optimize their performance on various kinds of likely multiprocessors. Secondly, we are systematically mapping the parallel processor parameter space to determine the most likely general features of future supercomputers. This, in turn, will help us to understand the structure of the algorithms that will perform most efficiently on those machines. Our critical needs in this area are sufficient resources to investigate fundamental issues such as algorithm decomposition, synchronization, and stochastic behavior.

Languages: The advent of parallel processors has a number of implications for high level languages (HLLs) and the compilers that translate them into executable machine instructions. These stem from increased logical complexity of parallel programs and the difficulties of managing parallel hardware resources in an efficient way. We are investigating an applicative language, VAL, which offers the potential for solving some of these problems. At the same time, LLNL has a large investment in Fortran software. Our strategy is to develop a modern procedural language for handling many code support functions, such as data base management, while providing Fortran language extensions and compilers for parallel processors. What will have to be done in the long run to programs written in FORTRAN (or similar languages) is a matter of current research. All of this work is basically exploratory in nature; to put these ideas into practice even on an experimental basis will require much more effort than is presently being expended.

Operating Systems: With the advent of parallel and distributed processors, operating systems are rapidly changing. As with HLLs, the challenges arise from increased logical complexity and stringent performance demands. We are actively involved in development of three different operating systems which address these issues in different ways. These are expansions of CTSS (Cray Time Sharing System) for the NMFECC, NLTSS (Network Livermore Time Sharing System) for the LCC network, and Amber for the S-1. These systems are all in advanced states of implementation, and they are putting us in a position to rapidly develop operating systems for future machines.

D. The S-1 Project

The S-1 Project is LLNL's principal area of effort in the supercomputer field and remains the only Government-sponsored (Navy/DOE) endeavor in high-performance arithmetic-intensive computing. The project is centered on the staged development of a set of design tools utilizing powerful, high-speed memory-intensive computers which support the rapid prototyping of general-purpose digital circuitry. When fully developed, these tools will support the ab initio design (with the order of a few man-years of effort) of a supercomputing system (including central processing unit, high-speed memory, I/O interfaces, diagnostics/maintenance unit, and operating system/utilities software) for implementation on a single semiconductor wafer.

These design tools, the S-1 Structured Computer-Aided Logic Design (SCALD) system, have been developed through most of three major stages, and have been successfully exercised at the completion of the first two stages to realize state-of-the-art computer systems. The third stage, when completed this year, will support the implementation of a state-of-the-art supercomputer on high-density printed circuit cards carrying VLSI gate arrays. The fourth stage is planned to be the final one, and is targeted on monolithic implementation of supercomputers.

The creation of a supercomputing system architecture that could be efficiently applied to a wide variety of problem areas, and that could grow with advancing implementation technology, has been one of the fundamental S-1 goals. The initial approach chosen was a MIMD interconnection between powerful computing engines and large memory units. Besides offering plug-in modular expansion capability, this approach places the least stress on development of new algorithms and software in order to exploit its latent capabilities.

Each of the computing engines of the MIMD system was designed to be a modern, general-purpose von Neumann machine compatible with multiprocessor requirements. This was done to minimize technical risk relative to algorithm development and breadth of applicability, while providing both systems and applications programmers with the set of features needed to maximize software engineering productivity.

A testbed for this new architecture was designed and built six years ago with the first SCALD stage. It features an overall computing performance comparable to IBM's then-top-of-the-line system, the 370/168, and carries a high-speed memory complement equal to the full memory capacity of such systems. This system, the S-1 Mark I, has also served to host the second-stage SCALD system.

The successor to this architectural testbed has been a much more powerful computing system embodying an extended version of the architecture proved out on the Mark I. The first of these second generation S-1 systems, designated Mark IIA, is presently completing check-out and software installation, and is being followed by five more Mark IIAs in the same packaging format. These six systems, each of which will be roughly comparable in performance with a Cray-1 system, will be interconnected into the first-stage S-1 Multiprocessor within a year. Re-packaging options aimed at minimizing cost, power consumption and size of these systems are presently being explored. We plan to produce enough re-packaged Mark IIAs (called Mark IIIs) so that, together with those already realized in the prototyping package, a total of sixteen will be interconnected into the S-1 Multiprocessor.

A minimal set of software just sufficient to exercise the capabilities of these hardware systems for test and evaluation purposes has been developed for the Mark IIA uniprocessors, as well as for the Multiprocessor into which they will interconnect. The widely used Fortran, Pascal, C and LISP languages are supported on the Mark IIA systems, which carry the Unix time-sharing operating system, as well as Amber, a Multics-like operating system aimed at fully exploiting the more novel features of these S-1 systems. Advanced software tools to perform the highly automated partitioning of programs for efficient execution on the S-1 Multiprocessor are awaiting evaluation on this system.

All aspects of the S-1 effort are continuing to operate at severely resource-constrained levels. At these levels the SCALD undertaking cannot seriously develop the capability needed to implement a monolithic supercomputer, the S-1 Mark V, in spite of the revolutionary gains in cost-performance ratio offered by attainment of this technological plateau. The implementation of the S-1 Multiprocessor has been stretched by budgetary stringencies to nearly twice the period intrinsic to this technological challenge, and the S-1 software efforts are similarly confined. Under these circumstances the project is in a position to considerably increase the rate of development over a short period of time (See Appendix D).

E. Component Technologies

The hardware component of any national supercomputing program needs the capability to acquire high-performance, high-density devices in adequate quantities, and with fast turnaround from device design specification to delivery of packaged and tested devices. Rapid progress in the S-1 program, for example, is closely tied to such a capability, particularly in its two future hardware generations, the Mark III and Mark V.

The computer industry clearly has or will install the capability to fulfill its interests. The rapid turnaround requirements of national laboratory efforts, such as the S-1 program, dictate a need for a captive state-of-the-art component facility. In addition, a major effort should be expended to form close working relationships between industry and any national laboratory computer design efforts to obtain the benefits of both environments.

LLNL's needs in this regard have already led to the creation of a modest microelectronics facility. In this facility, we develop unique components for our Weapons Test Program (both silicon and gallium arsenide), and we recently acquired a gate array design and implementation capability for rapid development of custom integrated circuits. We have also begun research at a modest level on a novel approach to high-speed, computer-controlled laser-centered "writing" of discretionarily placed and interconnected transistors on semiconductor wafers, a process that we call laser pantography. The different types of capabilities represented by these two approaches--the high volume and readily accessed features of gate array technology, and the promise of very high component count and arbitrary layout/interconnection aspects of the new laser pantography technology--are both likely to be essential to a national supercomputer program which contemplates any serious hardware implementations.

To meet these hardware goals, we propose that resources be provided to expand our capabilities in chip design, fabrication and testing, and to make these expanded capabilities accessible (e.g., via ARPANET and MILNET, nodes for both of which exist at LLNL) to other participants in a national supercomputer program. With such a new facility, we could significantly reduce the turnaround time for the high speed digital circuitry needed for future supercomputer projects. Typically, such research and development projects do not use sufficient quantities of custom-designed components to seriously interest commercial manufacturers, nor can the turnaround time be made short enough.

Initially, our approach will be to use commercially purchased gate array wafers (both silicon, ECL and gallium arsenide arrays when available) that we personalize with optical lithography, a capability that we have in limited measure today. Eventually, we would like to acquire a direct write-on-wafer electron beam lithography system. Concurrently, we will be investigating the new laser-writing techniques for defining and interconnecting all transistors directly on semiconductor wafers without masks. Once developed, these techniques can be used to build very small, high speed supercomputers using wafer scale integration with submicron feature sizes. The proposed microelectronics facilities (currently not funded) are discussed further in Appendix E, and the proposed laser pantography research (currently funded at a low level) in Appendix F.

IV. APPLICATIONS TO OTHER SCIENTIFIC AND DEFENSE PROGRAMS

As a guiding principle, LLNL's goal in scientific computing is to execute large scale modeling codes at the fastest possible rate with connection to the largest possible high-speed memory. In Section III we described our projected requirements in computing power, and indicated the actions we propose (and in some cases, are taking) to meet those needs. However, other constituents of the technical community have equally compelling but very different uses for supercomputers. For example, problems in which symbol processing is more frequent than arithmetic computation, or in which manipulation of large data bases is the main operation, may be more important for some defense applications. The architecture and languages that give optimal performance for these situations is likely to differ from those most appropriate for scientific calculations.

However, there are many points of common interest among the different generic classes of potential supercomputer users. Here we discuss how our experience, skills, and ongoing efforts can be usefully applied to such areas of general concern, and to specific problems of interest to these different communities.

A. Interactions with the National Computer Community

The national computer community is made up of university researchers, government laboratory staff members, defense scientists, and employees of private firms. Those who participate in a national supercomputer program should be able to communicate ideas and information freely and efficiently. This can be expedited through access to a common computer network and by coordinated dissemination of results in timely written reports and conferences. We believe that LLNL has a role to play in each area.

A national computer network can perform two functions. As in the case of the NMFECC (or possibly ARPANET) it can serve as a vehicle for research with supercomputers. It can simultaneously provide a mechanism for research on supercomputers. In an evolutionary program we envision a sequence in which the best of the experimental prototype computers evolve into the next generation main frames used for computing.

Our experience in developing and operating the NMFECC gives us the knowledge and skills to successfully operate a national network. Our motivation in taking on this task would be the opportunity to interact with a large community outside LLNL, and to host new prototype machines as they are conceived. Both of these factors would stimulate the supercomputing research effort within LLNL, and would justify the costs of using excellent operations people in what is partly a service role.

With regard to other means of communication, LLNL and the Los Alamos National Laboratory began hosting a series of conferences in 1981. These meetings were initially conceived as small gatherings (less than 100 people) in which the needs of the weapons laboratories could be communicated to computer manufacturers. They have evolved into more general forums in which the focus is on general questions of computing on parallel machines, and in which the participants span a broad spectrum of academic and industrial interests. A national supercomputer program would undoubtedly broaden the scope still further, but the relationships already developed should be very valuable in setting up an expanded agenda. As "neutral" government laboratories we think that both ourselves and LANL could continue to play useful roles in administering and focusing the aims of such conferences.

B. Defense and Systems Applications

It seems clear that within the next few decades, supercomputers will be developed with capabilities (including memory size) several orders of magnitude greater than those of today's leading edge machines. Extrapolation of current trends also makes it likely that those computers will be much smaller and much less expensive than present day supercomputers. Although the most important current objective is to formulate and implement a national program to bring those advances into being, it is also appropriate to explore the potential applications that might spring from such advances.

This will be a substantial exercise. There have been few technological advances in recorded history as explosive in character as that now ongoing in microelectronics and computation. The impact over the last 30 years has been profound; the impact over the next three decades is likely to be even greater. Even so, the rate of technology gains has been so rapid that the applications have lagged behind. Accordingly, we believe it timely to encourage an orderly conceptualization of such applications, and to consider mechanisms that could expedite this process.

As "mind stretcher" examples of the kind of thinking required, we have listed several recent LLNL reports in the Bibliography. The first paper (UCRL-86581) deals with the ultimate useful limits for "number-crunching" computers (as well as even more speculative matters in neurophysiology), and how they are likely to be attained. The second (UCID-19747) describes the computer requirements needed for real-time processing of synthetic aperture radar data. The third essay (UCID-19746) discusses the possible use of robotics in battlefield situations, and we describe its arguments in somewhat more detail.

Basically, the development of a battlefield robotics capability has two key requirements: target acquisition and weapon delivery. For target acquisition, estimates of the digital computational effort needed to inspect a high resolution image of a scene, and to recognize and precisely position objects of interest within it, suggest that it is within the current state-of-the-art of the underlying digital semiconductor technology, and that such a package might be accommodated within a large remotely piloted vehicle or tactical missile. As for weapon delivery, the paper also discusses the computing performance required for various levels of robot performance in military environments. Robotics experts estimate that minimal required performance levels are no greater than those required for real-time target localization and classification--of the order of 1000 MIPS of scalar processing.

What are the essential components of a program to achieve these levels of technological capabilities? They appear to be two in number. The first is a sustained, goal-oriented effort aimed at the early attainment of wafer-scale integration of very high performance computing systems. (See Appendix G) This is required for reasons of system radiation hardness, acceleration tolerance, and minimum cost, as well as minimum package size and mass. The second component is an effort aimed at developing robust algorithms for target recognition, localization, and classification in realistic military environments over the full class of targets of interest.

Although the numbers given in the essay are intended only as an illustrative example, they pinpoint the crucial "other" research component. Even if the supercomputing goals described in this paper are achieved, applications outside the number-crunching and signal processing arenas will require aggressive, concurrent development programs. We urge that such work be initiated now to fully exploit the anticipated technological gains in supercomputing.

V. BUDGET AND RESOURCE MATTERS

LLNL currently expends significant resources in using its major computers, and in maintaining modern operating systems and other software for those machines. Nearly all this work is dictated by day-to-day and near-term programmatic demands; only a small fraction is devoted to research and development on advanced supercomputers and their associated software requirements. In the latter context our current annual expenditures are approximately $1.0M of LLNL research monies (mostly Weapons Program) on advanced software issues, another $0.5M on mathematical studies funded by OBES, and about $6M on S-1 related projects (joint Navy/DOE sponsorship).

To meet fully the goals and proposals discussed in this paper would require the following incremental resources at LLNL:

• Computer prototype purchase and operation (example: multiprocessor VAX system, about $1.0M capital, $0.5M operating)	as required
• Software research and development - operating	$4-8M
• Accelerated/expanded S-1 Project - operating	$6-9M
• State-of-the-Art Microelectronics Facilities - capital	$10M

The hardware efforts, although resource limited at present, have the necessary high quality staff in place. The software program would require a significant expansion of the research base, and integration of those activities with the ongoing programmatic work. Some of the capital items depend on specific role assignment and commercial computer availability; the flexibility for rapid acquisition will obviously be very important.

VI. APPENDICES

This series of Appendices provides expanded descriptions of most of the topics discussed in Sections III and IV.

Appendix A. Computational Needs of the Nuclear Weapons Program

This topic is discussed in a classified report provided separately. Its document number is CND 83-13.

Appendix B. Computers at Livermore: History, Status, and Plans

Since the early 1950s, LLNL has acquired the most powerful computers available in order to execute the large physics modelling codes necessary to design nuclear weapons. As a result, many times over the past 30 years LLNL has purchased the first or early serial number of a new line of computing equipment:

Machine	Serial No.	Arrival Date
UNIVAC-1	2 or 3	March 1953
IBM-701	early	March 1954
IBM-704	early	March 1956
LARC	1	September 1960
IBM-7090	1	July 1960
IBM-7030	2 or 3	July 1961
CDC-1604	1	January 1963
IBM-7094	1	February 1963
CDC-6600	1	October 1964
CDC-7600	1	January 1969
CDC-STAR 100	1	November 1974
CRAY-1	6	May 1978

It is often the case that the first few models of a new computer line to reach LLNL are devoid of usable software for the numerical modeling community. In the case of many of the above machine types, LLNL found it necessary to develop and provide the total operating environment plus all the software in order to integrate the computer into the local computing center. This effort included, besides the application codes, the operating system, high-level programming languages, interfaces for input/output and data storage.

During the mid-1960s LLNL adopted the practice of timesharing its largest computers in a network configuration, a practice that is continued today. The Livermore Computer Center (LCC) currently provides a time-shared network access of 4 CRAY-1s and 3 CDC 7600s to the 2500 users at LLNL. Integration of a new type of computer into this network is a complex process requiring many man-years of effort.

Consequently, LLNL has the experienced personnel to facilitate the introduction of software-poor "number crunchers" into the existing environment. The network is a rich environment for storage and I/O as well, since approximately 3.3×10^{12} bits are on-line to the major computers, and we produce about 13 million page equivalents of output each month (film plus paper).

In addition to the LCC, LLNL developed and operates the National Magnetic Fusion Energy Computer Center (NMFECC). This national network, centered on two CRAY-1s and a CDC 7600, services over 40 sites around the U.S. that are involved with research in magnetic fusion. This network uses commercial satellite communications for data transfer to/from its major sites.

It is also appropriate to note that our LCC timesharing system has been adopted by NMFECC, the Los Alamos National Laboratory, and the Air Force Weapons Laboratory. In the case of NMFECC this system was substantially modified to provide more flexibility and dynamic file capabilities.

As the next generation ("class") of computer develops in the 1980s, LLNL will want to be funded to acquire the best available "supercomputers". Since these "Class VII machines" will likely be some form of multiple processor architecture, it is very important that initial research and advanced development work be done to evaluate the various potential configurations. Research will be necessary to effectively utilize these new machines, and thereby keep to a minimum the man-years of conversion that will be necessary to adapt application codes to a multiple processor. Unless we can obtain the resources to carry out this fundamental research expeditiously, converting to a CLASS VII computer in our environment would be more costly in effort than any of LLNLs transitions to previous computer acquisitions.

Since hands on experience is a valuable way to study the characteristics of a machine and measure its effective performance level, we believe that a national supercomputer program should have as an essential part the purchase and placement of state-of-the-art (prototype) computer systems for use by researchers. LLNL is in an excellent position to offer the proper environment for hosting some of these prototype supercomputer architectures. We have the internal "need," the applications, and the "people" who can collaborate with others from academic organizations and industry.

Within the next 3-4 years, the following six multiple processor architectures seem likely candidates for "hosted machines":

- Control Data Corp: "Cyber 2XX"
- Cray Research Inc: "CRAY-II"
- Denelcor: "HEP"
- LLNL: "S-1 (Mark IIA)"
- Control Data Corp: "AFP"
- Digital Equipment Corp: "MP VAX" (4-6 PDP VAX 11/780s)

Each of the six listed machines offers a different approach to packaging several processors into a single computing resource. Except for the "MP VAX," all of the above machines are likely to be "software poor" when initially available to the national market place. Given the resources to attain and support one or more of these machines, the following activities and issues are likely to be pursued and the results made available to the national research community:

- Hardware benchmarking and evaluation
- Algorithm testing
- Programming language experimentation
- Operating systems development
- Storage requirement determinations
- User "environmental" studies

Our current goal would be to host several prototype machines. The S-1 is a possible choice because of propinquity and the ability to carry out tightly coupled hardware-software development. As a national host we would also undertake the obligation to supply other network users with an accessible and friendly computing environment.

We have made pilot studies of the feasibility of hosting the other machines. The particular choice would depend on the available resources, the operational readiness of the machine, and the near-term goals of the associated software program.

Appendix C. Software Research for Supercomputing

The transition from serial to vector computers during the 1970's was much more difficult than anticipated by manufacturers, computer scientists, or users. Two reasons for this stand out. First, although the performance of the vector machines was excellent on special kinds of problems, most existing codes ran slower, or marginally faster, than on the previous generation of serial computers. Secondly, modification of those codes to take optimal advantage of vector capabilities was an arduous and complex task. The structure of the existing programming languages and application codes was simply ill-suited for adaptation to vector processing. In short the manufacturers/computer scientists ignored the character of the application codes, and the users/programmers ignored the potential of vector machines. The result was a several year period of no real growth in realizable computing power for large scale scientific and engineering codes. However, it was also a process in which we learned a great deal about the kinds of software issues that accompany a major shift in computer architecture.

The effort in going from vector to parallel computers may be even more extensive and complex than the serial to vector transition. But, this move to parallel machines also seems inevitable if our ability to compute is to keep pace with the demands of the users. Consequently, we have initiated software projects in several areas aimed at bridging the gap. These are

- Simulators for Multiprocessors
- Parallel Algorithms
- Modular Code Systems
- High Level Languages
- Numerical Methods

Although each of these studies has produced interesting preliminary results, a critical mass for significant progress will require several-fold increases in resources over current levels (couple of people per year per topic). In addition, of course, there are myriad software projects associated with efficiently using any new computer. These include operating system development, editors, innovative graphics, and the host of other features that contribute to a productive user interface with the computer. At present our activities in these areas are largely restricted to programmatic work designed to provide near-term receptiveness for the next production supercomputer acquired by LLNL.

An outline of the overall software research structure is shown in Fig. C-1. In the remainder of this Appendix we concentrate on those projects that are specifically designed to handle the parallelism and other innovations inherent in advanced new computers.

Simulators for Multiprocessors

For all the reasons discussed above, we see a critical need to understand the performance and characteristics of parallel computers. However, the limited availability of multiprocessors, particularly ones with usable software for implementing algorithms, makes an approach which relies exclusively on existing hardware take unacceptably long. Consequently, our approach is to test algorithms on actual computers as they become available, and additionally to develop the capability of simulating MIMD (Multiple Instruction/Multiple Data) machines. The simulation capability will be useful both for near-term analysis of multiprocessors such as the S-1, and for more general studies of multiprocessor algorithm/machine performance.

In our early work we have focused on multiprocessors consisting of a relatively small number (typically no more than 16) of identical computers.

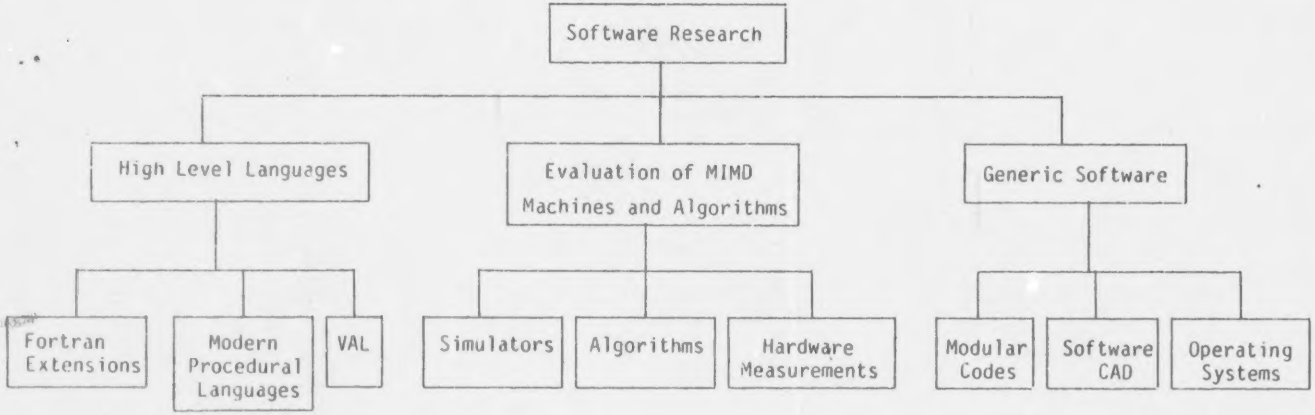

Software Research
High Level Languages
Evaluation of MIMD Machines and Algorithms
Generic Software
Fortran Extensions
Modern Procedural Languages
VAL
Simulators
Algorithms
Hardware Measurements
Modular Codes
Software CAD
Operating Systems

Fig. C-1. Software Research Structure

Each computer can execute a program independently of the others, but in general will be cooperating to perform a single calculation (the S-1, Cray X-MP, and Cray-2 all fit this description). In the past year we have applied our MIMD simulator to the S-1 Mark IIA multiprocessor and investigated the behavior of several parallel algorithms. In the future we plan to validate our simulator results with actual measurements on the S-1 Mark IIA multiprocessor, when it becomes available, and to study other potential MIMD machines if additional resources become available.

Parallel Algorithms

The highly parallel nature of advanced computers will require algorithms for numerical simulation to be significantly redesigned to use large numbers of cooperating and communicating parallel processes. Loss of efficiency due to frequent synchronization or interprocess communication must be minimized. The technical issues that must be addressed are

- Algorithm decomposition
- Synchronization costs
- Stochastic behavior
- Memory and switch conflicts.

Our approach to these problems is twofold. We have first constructed an initial set of benchmark codes appropriate to LLNL application: a 2-dimensional hydrodynamic/transport code, a Monte Carlo program, and a biconjugate gradient linear solver. These have been debugged on the MIMD simulators, and we have begun testing their performance on the multiprocessors as represented by the simulators. This has begun to provide the first extension of the early Carnegie-Mellon results on speed-up as a function of the number of processors.

Our second goal is to systematically map the parallel processor parameter space, using analytic tools or simple queing models to estimate performance. This is useful because it identifies parameters to be measured in actual machines, and it gives some advance indication of how computers will evolve. When these results are combined with the trends in hardware--from cycle time to memory characteristics--we can work backwards to define the properties of likely supercomputers. This in turn helps us to understand the structure of the parallel algorithms that will be needed to optimally use such machines.

In both the above areas, our studies to date have provided qualitative insights on scientific computing in a highly parallel environment. They have also pointed out key issues--such as stochastic behavior--in which major research needs to be pursued. Further advances in these pivotal areas will require either additional resources or restricting our attention to only one or two problems.

Modular Code Systems

Numerical modeling codes, and probably most software systems, are increasingly required to be highly adaptable, so that new advances in modeling techniques can be rapidly incorporated or the code generalized to application areas outside its original domain. This adaptability can only be attained through a major increase in the modularity of software systems. The rapid transition to computers with high degrees of parallelism also has implications for the overall structure of applications codes. It is not sufficient for the algorithms to be expressed as large numbers of cooperating parallel processes, although that is certainly necessary. The infrastructure of the code, which handles database functions, interprocess communication, and input/output, must be similarly structured. In our current codes, these functions account for roughly 75% of the total lines-of-code, and their redesign is an important and early task.

We have initiated a program to address the problems associated with the development, management, maintenance, and operation of the computer software needed for large scale simulation of nuclear weapons on advanced computers. This project, named MERLIN, is developing a system of code modules which perform highly specialized but frequently required functions. Many of the modules we hope to develop would be transferable to other research facilities. The modules would control processes such as editing, graphics, problem generation, language interpretation, data base management, memory management, and various algorithms necessary for numerical simulation of nuclear weapons.

This system of modules will be controlled and run by a sophisticated management code.

At present much of our attention has been focused on the needs of 2-dimensional hydrodynamics codes, specifically on automatic linkers, rezoners, plotters, and other grid-to-grid service functions. The pace of our development is basically set by resource constraints: to achieve the relatively general modular management system within a few years will require a two- to three-fold increase over our current levels.

We are also developing software subsystems which provide numerical mathematics services to modular physics codes. The emphasis in this work is to create systems which have a high degree of autonomy within their own domain and present a flexible and well defined interface to the external world. Our work is most advanced in the area of large systems of ordinary differential equations (ODE). We are expanding this work to include modules which provide interfaces between finite element or multigrid codes and ODE solvers.

High Level Languages

The advent of parallel processors has a number of implications for high level languages and the compilers that translate them into executable machine instructions. These stem from the increased logical complexity of parallel programs, and from the difficulties of managing parallel hardware resources in an efficient way. We are investigating an applicative language, VAL (Value-Oriented Algorithmic Language), which offers the potential for solving some of these problems. However, LLNL has a very large investment in Fortran software (over two million lines of coding) that cannot practically be abandoned. Consequently, we plan to provide Fortran language extensions for parallel processing and continue our development of advanced compilers for it.

As noted previously, the majority of the lines of code of an applications code system deal with the "support" of the calculation, and are not computationally intensive. This portion of the code system is, however, very complicated with respect to control logic, data structures, exception handling, and high level process synchronization. It is, in fact, a microcosm of an operating system and Fortran has gradually become outdated in many significant areas as a language for this environment. Therefore, we are exploring the development of modern procedural language which will eventually replace Fortran as the languages of choice for those codes/modules which support the numerically intensive physics calculations.

Our work on the applicative language VAL has grown out of a collaboration with MIT (the initial definition of VAL was published in 1979). The language is very attractive for us because it is designed for computers capable of highly concurrent operation. Some of the specific benefits are:

- Guaranteed deterministic results
- Deadlock-free
- Implicit task division and assignment
- No explicit memory allocation
- Implicit synchronization
- No side effects

We are currently working on VAL compiler development for three different target machines: The VAX-11/780, Cray-1, and the HEP-1. Our strategy is to get the compilers working, begin running test problems, and start addressing the central issues in parallel computation. The early experience with VAL has already demonstrated some of the benefits of an applicative language--the determinism, lack of side effects, etc.--but it will require major effort to solve the partitioning and memory use problems.

We also propose to make further extensions of the language to include vectors, pipelining, multiple functional units, and multiple processors. As new machines become available we hope to proceed to a compiler-code generator sequence for each computer. As we gain experience with real multiprocessors we could thus be able to further improve VAL for the kinds of large scale computations that are most important for scientific applications at LLNL. These latter efforts are largely beyond our own resource base. We have partially alleviated this problem through cooperate work with universities (from MIT to Manchester). However, additional funding for these ventures is imperative for major progress.

Advanced Numerical Methods

For the past eight years, the Applied Mathematical Sciences Division within the Department of Energy's Office of Basic Energy Sciences (OBES) has sponsored a small basic research program at LLNL on advanced numerical methods for large scale scientific computing. Originally, this program concentrated on vector processors. Calculations performed on vector processors (the most advanced computer at that time) were not able to achieve maximal computer performance because appropriate numerical techniques were not known. The OBES-sponsored work established a theoretical knowledge base from which many new vector algorithms were developed.

In 1979, it became apparent to the OBES that an adequate knowledge base in large scale scientific computing had not been established to guide the development of even higher performance computing systems. A more global approach had to be taken, whereby the interplay of the mathematical model, the numerical approximations, the numerical algorithms, and the computational hardware had to be studied. Accordingly, the OBES expanded its program to include contributions from the fields of computational physics, computational mathematics, numerical analysis, and numerical computer science.

This expanded program is divided into three efforts: the creation and collection of representative benchmark computational models; the application of mathematical research to the solution of numerical partial differential equations; and the performance of research in numerical computer science for large-scale scientific computing. In the first effort, an initial collection of computational models from the fields of hydrodynamics, magneto-hydrodynamics, transport, and molecular dynamics has been accomplished. The intent of these models is to provide general statements of computational and mathematical problems so as to guide, in a uniform manner, the other two OBES efforts. A goal of this effort is to provide a document stating the mathematics of the model, as well as its performance on various advanced computing architectures, so as to reveal open questions of research to computational scientists in and outside of DoE.

The motivation for the effort in numerical partial differential equations is the following: In order to re-organize the numerical methods being used for the solution of partial differential equations on an advanced computing architecture, it is necessary to first establish a knowledge base so that the flexibility of these methods can be determined. The areas currently undergoing research in the OBES program fall into the categories of numerical approximation for time-dependent partial differential equations and numerical algorithms for large sparse linear systems. In the category of numerical approximations, research is being conducted on wave tracking methods and on the effects of local grid refinement and moving node methods on wave propagation problems. Such problems are of fundamental importance in the fields of hydrodynamics and magnetohydrodynamics. Also being studied are asymptotic methods for the resolution of boundary layers and turning points in advection/diffusion equations. In the area of numerical algorithms, research on multi-grid methods for solving the implicit system of equations arising in diffusion models is being carried on. All of these areas are of basic importance to the construction of efficient and accurate large scale numerical models.

Work in numerical computer science falls into two categories: parallel numerical methods and high-level scientific languages. In the area of parallel methods, partitioning techniques are being studied in both ~~the~~ spatial and temporal domain for time-dependent partial differential equations. In high level languages, research is being conducted to determine the weaknesses of the Fortran language in expressing the newer adaptive methods with respect to problems in numerical hydrodynamics.

A primary need in all of the numerical methods work is to understand its implications for architectures and languages, and to integrate its specific results into less specialized applications codes. Bridging these gaps is a major overall objective of our software program, and an area in which we hope to see an increased commitment of resources.

Appendix D. The S-1 Project

Background

Very high bandwidth processing of data has been recognized as being a pacing technology in many areas of National defense; one of the more recent reports stressing this point was that of the Defense Science Board's 1981 Summer Study. Both acoustic and non-acoustic ASW, most of the varieties of orbital reconnaissance, ballistic missile defense, and naval engagement and battlefield management are representative of the many areas of defense technology which critically rely on the capabilities of available data processing means.

Digital data processing is widely recognized as offering great flexibility in data processing. For some national security applications, such as those of NSA and DoE's Defense Programs, digital means of manipulating data seem absolutely essential. It is therefore not surprising that key aspects of the digital technologies have been vigorously fostered by these portions of the national security community since technical feasibility was established four decades ago.

Underlying the widely perceived prospects for digital data processing to strongly influence national defense during the next two decades has been the remarkable development of semiconductor digital electronics during the past decade and a half. The figures-of-merit of digital electronics--switching speed, device density, device unit cost, switching speed-power product--have each advanced by several orders-of-magnitude within a fraction of a human lifetime, a characteristic of only the most revolutionary of human technologies. Moreover five to six orders of magnitude in devices per substrate and in device switching speed-power product remain to be gained before physical limits are reached, at least as much technological territory as has been covered since the invention of the transistor.

At the present time, microprocessors and their specialized relatives are creeping in at the fringes of defense technology, e.g., in the VHSIC Program, but with generally quite modest performance goals for most of a decade hence, relative to existing supercomputers. On the other hand, the technological fallout in military areas deriving from attempts to advance supercomputing capabilities by a thousand-fold, is likely to be very large. Since faster computers are necessarily smaller, lower power consumption ones, they are more readily adaptable to use with relatively low cost, mobile systems. "Effectively sentient" military robots capable of at-least-human performance for a limited set of tasks (e.g., fighting a tank, piloting a remotely piloted vehicle, navigating a tactical sub) thus appear to be likely fallout from vigorous supercomputer efforts.

It was with such considerations in mind that the S-1 Project commenced investigation of the technological bounds on defense-oriented supercomputing system performance in 1975. It successfully sought Navy sponsorship for these initial efforts, due to the Navy's pre-eminence within the DOD community with respect to computer utilization for weapons systems control. It has worked since that time primarily under Navy sponsorship, with recent co-sponsorship from the Department of Energy's Office of Military Application. Its total funding to date has been 18 million of then-year dollars, and its FY-83 budget guidance from its Navy and DOE co-sponsors totals over $6 million.

Technical Approach

The S-1 approach to supercomputing is premised on the basic lesson from LSI/VLSI--it's currently faster, easier, and cheaper to build ten computers than it is to build one computer which runs ten times faster--constrained by the distinctly finite amount of parallelism present in most defense related problems. The criticism of a single extremely fast computer is that it is usually extremely expensive to implement, and thus of limited replicability and limited (though positive definite) applicability to defense problems. The S-1 view of the million-microprocessors-tied-together approach to very high performance computing is that "effective" interconnection of this many computing elements is difficult to realize and even more challenging to exploit with high-level software. Furthermore, extraction of sufficient parallelism to make efficient use of such an ensemble for a reasonably wide spectrum of defense problems is technically very risky, particularly because so little progress has been made in pertinent simulation studies and exploration with small microprocessor networks during the past decade.

Supercomputing in the S-1 context (for at least the next half-decade) is therefore based on the creation of a substantial but not large number (e.g., 16-256) of powerful, von Neumann-esque digital processors, interconnected to a common, very high capacity semiconductor memory system, as well as to each other, via a high aggregate bandwidth full crossbar (crosspoint) switch. The total system, a S-1 multiprocessor, is thus a Multiple-Instruction Stream/Multiple-Data Stream (MIMD) one, with its member processors able to work completely independently of each other on one or more tasks each, in partial to complete collaboration on a number of tasks, or in maximum collaboration on a single task. Each of the member processors has large, separate data and instruction cache memories embedded in it to maximize bandwidth to its instruction preparation and execution units and to minimize required traffic through the crossbar switch. Parallelism appears within such processors in the now-standard modes for high performance pro- cessors: pipelined preparation and execution of instructions, including vector instruction execution. Unique instruction preparation and execution features created specifically to address current defense computing needs are included in each processor (e.g., very efficient emulation of existing DoD computers, and ex- tremely high speed execution as single hardware-implemented instructions of the most useful scientific subroutines). From architecture design to implementation, the pacing role of RAMs in the evolution of the underlying digital electronics technologies is affirmatively acknowledged: S-1 CPUs and their associated high speed memory units use RAMs of all performance levels very aggressively to attain cost-efficient performance.

The present (second) generation of S-1 processors, the Mark IIAs, are each expected to compute between one-third and ten times as fast as a CRAY-1 performing equivalent functions, depending on the type of code being executed. Simulation studies have suggested that a typical mix of scientific processing will proceed between 9 and 15 times as rapidly on the 16 member S-1 Multiprocessor as it will on a single Mark IIA uniprocessor.

Central to the creation of S-1 processors, both past and future, and of major value in its own right, is the S-1 Structured Computer-Aided Logic Design (SCALD) system. This body of nearly 100,000 lines of Pascal implemented software and its supporting hardware serves as a compiler-and code-generator for general-purpose digital hardware, complete with both source and object time debugging support. It provides digital logic designers with a quite complete, extensively computer-assisted 'pathway to silicon' from high-level digital architectural concepts, extensively supports post-implementation debugging, and generates comprehensive documentation of the entire process. In its most recent complete form, SCALD II, it is targeted to implementation on wire-wrap boards populated with commercially multi-sourced MSI and LSI circuit packages and interconnected with flat cables. SCALD III, presently under development, aims at system implementation onto printed circuit cards and gate arrays, while SCALD V, currently under study, is directed to monolithic implementation of very high performance digital systems.

The creation of supercomputers hitherto has been slowed by the paucity of both computing power and high-speed memory capacity employed in the design and implementation processes. When employed for supercomputer design, SCALD exploits the unusually large high-speed memory systems carried by S-1 computers and helps break this bottleneck. Indeed, the present approach to realizing the S-1 Mark III and Mark V uniprocessors requires the huge virtual and real memory capacities of the S-1 Mark IIA processors, exercised by the very large computing bandwidth of these systems, as these Mark IIA systems previously depended critically on the availability of the S-1 Mark I system to run SCALD during their design and implementation.

To support sponsor evaluation of the supercomputer systems which it has been creating while developing and exercising SCALD, the S-1 Project has transported, as well as created ab initio, a set of software environments emphasizing the use of high-level programming languages and ultimately intended to feature the full set of modern software engineering tools. This work has been greatly aided by the basic architectural upwards-compatibility of each S-1 hardware generation with respect to the previous ones, thus facilitating software transport from earlier S-1 systems, as well as from non-S-1 systems. To date, these efforts have produced Fortran, C and Pascal compilers and associated run-time environments, an interpreter and an S-1 architecture-targeted compiler for the NIL dialect of LISP, Version 7 of tne Unix operating system, and the basic features of Amber, a Multics-flavored operating system developed to express and exploit the full hardware capabilities of both S-1 uni- and multi-processors. The most popular utilities and system services (e.g., compilation and phototypsetting of technical documents) of an upper-level ARPANET host are also available on the S-1 local network, in addition to the special system features which support intensive SCALD usage.

Recent Progress

The S-1 Project received the funding required to commence implementation of an S-1 Multiprocessor system in second generation (Mark IIA) hardware in the final quarter of FY 81, and issued implementation instructions for a pair of Mark IIA uniprocessor prototypes to its supporting sub-contractors in the first quarter of FY 82. (These instructions were generated by SCALD II running on the S-1 Mark I processor--roughly equivalent to 4 DEC KL-10s in processing power--equipped with 16 megabytes of high-speed memory, and consisted of 200 megabytes, including SCALD-generated design and implementation documentation.) Most of FY 82 was devoted to manufacturing and high-level integration of these two prototypes, and to early phase construction of the following four Mark IIA uniprocessors (Serial 3, 4, 5, and 6). Debugging of the Serial 1 uniprocessor commenced in late 1982, and has progressed to the point of successfully running several dozen small test programs at full clock speed. The Serial 2 machine has been following the Serial 1 Mark IIA with respect to removal of design and implementation flaws, and is expected to be available within a quarter of commissioning of the Serial 1 system, reflecting 'learning curve' gains. At least four more Mark IIA systems will then be implemented, and all Mark IIA systems interconnected into the S-1 multiprocessor.

SCALD III, a major extension of SCALD II emphasizing fully automatically generated implementations in high-density printed circuit boards and LSI/VLSI gate arrays, had most of its printed circuit optimizing routing package developed and debugged during FY 82. This system is presently being validated in the course of creating the large (16 megabit) memory cards which will populate the Memory Boxes of the Mark IIA Multiprocessor. During 1982, SCALD II software and documentation was distributed under Navy supervision to approximately 150 requestors of it in American industry, academia and Government. Among these recipients were two companies (Daisy and Valid Logic Systems) which are already quite successful, with SCALD-based CAD/CAM systems as their products.

The 1982 period also saw completion of the Base System of the Amber uni-/multi-processor operating system created to exploit fully S-1 system features, and its exercising on the Mark IIA architecture simulator; a comparable milestone was also attained for the Unix (Version 7) uniprocessor operating system being transported into the S-1 architecture from that of the DEC VAX. A complete NIL/LISP system has been extensively exercised on the S-1 architecture simulator during 1982. 'Cold load' versions of Amber, Unix and the LISP environment have been prepared for installation and exercising as soon as a Mark IIA system becomes available for software development. The "universal" code generator/optimizer for the S-1 Fortran and Pascal compilers was upgraded during FY 82 to target the full Mark IIA architecture, and the Symbolic Debugger common to all S-1 high-level languages has been fully commissioned. Work was also commenced during this past year on a Fortran compiler aimed at exploiting the S-1 Mark IIA vector instruction set, and the extensive hardware implementation of this major portion of the S-1 Native Mode architecture; however, present resource constraints preclude completion of this work in time to make the vector capabilities of the Mark IIA available from a high-level language during the next two years.

Planned and Proposed Work

Co-sponsorship of the S-1 Project by the Navy and the DOE during FY 83 has made available a total of over $6 million, roughly two-thirds of which is required for Project manpower and housekeeping costs (e.g., leasing of office/lab/Multiprocessor Facility trailers, ARPANET IMP connection charges, etc.), and approximately one-third of which is to be devoted to completion of the first six Mark IIA uniprocessors, and 6 Memory Boxes containing a total of a quarter-gigabyte of semiconductor RAMs.

Because of the high 'ante-in' costs already in place and representing such a major fraction of the present resource envelope, the pace of R&D contemplated with $6 M of funding is considerably below (about a factor of two) that which could be undertaken without undue programmatic risk. Because the Project's staff and facilities are already 'in place,' more rapid completion of the full-scale S-1 Multiprocessor (planned to be constituted of 16 uniprocessors and 16 memory units), earlier development of the S-1 Mark III uniprocessor, fuller development of major S-1 software modules (e.g., the upper layers of the Amber operating system, NIL/LISP environment, and ARPANET access) and more aggressive development of the monolithic implementation oriented SCALD V all could be executed with high return on incremental investment. Most additional funds committed to hardware developments would be devoted to component and subsystem purchases, rather than to manpower costs, while enhanced support for S-1 software development, though devoted primarily to defraying additional manpower and academic sub-contract costs, would leverage off the significant investment in software development tools and programmer support systems already present in the S-1 local computing network.

It is therefore proposed to accelerate the pace of S-1 developments in each of these major endeavor areas with additional funds. Such acceleration would move the S-1 supercomputing effort toward its next set of milestones at a much more rapid pace, significantly augmenting the technical and programmatic momentum that has been built up over the past seven years in the Nation's only (until very recently) Government-sponsored supercomputing project. It would serve to naturally integrate into a national supercomputing program the capabilities and experience of the S-1 Project pertinent to this program's nascent efforts. It would provide a natural institutional vehicle for the research community's participation in evaluation and exploitation of the S-1 Multiprocessor. Such real experience on such a large-scale, high performance-and-memory capacity machine should be of great value, both in learning how to best use such items and on determining what further optimizations are desirable. For example, S-1 Mark IIA uniprocessor systems will quite certainly be the highest throughput, largest memory capacity LISP engines for some time, and the S-1 Multiprocessor will provide the first large-scale opportunity to explore LISP-expressed problems in a multiprocessing environment.

Appendix E. Expanded Microelectronics Facilities

A national supercomputer program will need many types of high-performance, high-density integrated circuits. Many of these will be custom built uniquely for supercomputer applications. For example, the S-1 Mark III computer planned for the next phase of the S-1 supercomputer project will use 50 to 200 different personalizations (designs) of ECL gate arrays. In addition, several promising proposals for supercomputers involve new integrated circuit technologies, such as wafer scale integration, which require continued research and development efforts and which will not be available from commercial sources in the foreseeable future.

To meet these needs we believe that a state-of-the-art microelectronics facility is required. A captive, modern facility at LLNL would permit rapid development of adequate quantities of the custom circuits needed by the various hardware design projects of a supercomputer program. Manufacturers will not generally produce, in a timely manner small quantities of custom circuits, because of the relatively low return on investment relative to commercial customers who have annual production needs. Also, by supporting a facility and staff whose central focus is the development of circuits for this program, there will be greater coupling of the ideas between system architects, hardware designers and semiconductor technologists.

A microelectronics facility for a national supercomputer program will need to develop and produce a succession of new parts, whose level of complexity and performance increases with succeeding milestones of the overall program. We believe these goals could be met by a progression of new manufacturing methods for producing more dense custom circuits, and possibly by development of new material and heterojunction transistor designs to challenge conventional silicon transistors in some low density applications. The approaches for customizing circuits will change from gate array technology to more dense full custom integrated circuits, and finally to wafer scale integration. Today, the most useful gate array technology is silicon ECL. In the future, higher speed gallium arsenide MESFETS and III-V heterojunction bipolar gate arrays may become available since several commercial efforts in this area are underway. For the more dense full custom and wafer scale integration approaches, the most useful transistor technology will be silicon MOS, with perhaps some ECL bipolar for very fast I/O functions.

Our present LLNL facility enables us to produce the gate arrays needed for the early phases of the S-1 supercomputer project. For the next phase, the Mark III, the design and implementation of the 50-200 different personalized ECL gate arrays will require a $0.2M upgrade of our laboratory facilities, primarily by adding clean room capabilities. This upgrade is underway, and should be completed in the next year.

Implementation of the Mark V, however, would require development of a significantly expanded microelectronics facility. Technically, it will need wafer-scale integration techniques using logic gates with about 100 ps to 500 ps loaded gate delays and memory technology capable of a few nanoseconds access time. To achieve this, it now appears that we can combine in a single silicon wafer MOS technology with .5 micron channels and silicon ECL with 1 to 2 micron channels. To develop this complex system on one wafer we would use redundant designs in the system architecture, and would use new manufacturing techniques that will greatly improve the overall fabrication yield. The key to improving the yield is to eliminate manually introduced defects by automating the processing steps, and to use manufacturing techniques that will eliminate all conventional masking steps.

Even with such improvements, it is not clear that Mark V requirements can be met optimally by these technologies. If not it may be appropriate to look to III-V bipolar heterojunction technology (such as GaAs) under development in industry and at LLNL. To carry out research in these III-V compound transistors, a molecular beam epitaxy capability is needed in addition to the other fabrication facilities required for silicon circuits.

To achieve these Mark V technological goals, we require a substantially larger laboratory that has precise regulation of the atmospheric and seismic environment. The major categories of equipment needed in the facility include equipment for submicron lithography; growth of thin films of metals, dielectrics, and semiconducting; surface analysis; and a variety of precision diagnostic tools. The total capital cost for the new facilities would be on the order of $10M (no funding has yet been realized).

Such a facility could be accessible (e.g., via ARPANET and MILNET) to other participants in a national supercomputer program. LLNL already has significant contacts with industrial and university researchers who can provide essential skills. We have begun discussion with UC Berkeley, UC Santa Barbara, and UC Davis faculty to develop mechanisms for their participation. UC Berkeley's recognized expertise in all aspects of computer-aided circuit design would greatly improve those aspects of the program. UC Santa Barbara specializes in III-V semiconductor device and materials research, and could contribute much to the development of new fast logic based on III-V heterojunction transistors. We already have a small contract with UC Davis concerned with CMOS gate arrays. We plan that researchers from these and other universities, and from industry, would have access to the facility for research projects concerned with the supercomputer program goals.

Appendix F. Laser Pantography

Ultra-large scale integration (ULSI) of top-of-the-line computer systems (CPU, high-speed memory, system software, I/O interfaces, and diagnostics and maintenance subsystems) onto a single semiconducting substrate is highly desirable for reasons of attainable circuit speed, economy, reliability, and compactness. The present barriers to realization of such ULSI circuits are those of design complexity and fallibiility, implementation imperfections, latencies, and costs.

Much of LLNL's S-1 future program consists of redesigning the S-1 supercomputer down onto mostly CMOS digital logic and memory (instead of the ECL circuitry used in the first and second generation S-1 systems), and packaging this circuitry onto a single semiconductor substrate. The nominal approach to realizing such ULSI computer systems is via silicon foundry service, which would take in mask-specifying tapes written by the fifth generation of the SCALD system, SCALD V, and produce prototype wafers. However, in 1982, the few silicon foundries which exist have quite large (half-million dollar scale) front-end charges, substantial ($30-75 K) turn-around costs and, most dauntingly, several month turn-around times. They are clearly oriented to producing very many copies of small, high reliability designs, not for prototyping ultra-large scale circuits with requirements for many design iterations.

Whether silicon foundry service will evolve sufficiently rapidly in the directions required to effectively support ULSI prototyping during the next few years is quite uncertain, and constitutes one of the major sources of technical risk for the later phases of the S-1 program. A conventional alternative would be an effectively dedicated wafer processing line, whose set-up and operation cost for the required period has been estimated at about $10M. (See Appendix E)

However, it is technically feasible to directly "write" patterns of diffusion, dielectric, and metal--e.g. MOS digital circuitry--on a semiconductor surface, as has been demonstrated by several groups over the past two years. Heating of a surface with a focussed, pulsed laser beam over which an atmosphere of suitable composition and pressure is maintained has been shown to result in p-n junctions, dielectric sheets and metal lines on silicon and other semiconductor surfaces, depending on the composition of the gas in which the surface is heated. We call this process laser pantography. The dimensions and qualities of these modified regions are suitable for use in very high performance MOS devices. The rates at which devices can be created and interconnected appear to permit realization of at least 100,000 transistors per second; the 40 million transistors of the S-1 Mark IIA CPU would be produced and connected at these rates in less than a half-day.

LLNL has had a small effort studying the possibilities of fully automated creation of ULSI circuitry with these techniques for two years. It has recently demonstrated the high-speed, completely computer-controlled creation of micron-scale metal lines of arbitrary (planar) geometry of qualities suitable for interconnecting devices in a ULSI circuit. It demonstrated the ability to use a focussed laser beam to selectively mill silicon structures earlier this year (e.g. to remove unwanted material from circuit elements). It is presently proceeding to create and characterize micron-scale MOS transistors fabricated with this technology. Upgrading of the experimental capabilities now underway will result in complete computer control of all aspects of wafer positioning, laser beam scanning, focussing and amplitude switching, and gas composition and control, in addition to computer interfacing all real-time diagnostics. These diagnostics now include optical and will soon include electrical monitoring of the structures being constructed on the semiconductor surface.

If suitable resources were available this effort could proceed to synthesize small-scale integrated MOS circuits during this year, a "standard family" of MSI circuits and RAMS the following year, and commence ULSI circuit synthesis studies in the third year (using SCALD V capability from the S-1 Project which should be available at that time). The estimated costs for this work would be approximately $1.0M, $1.3M, and $1.7M in successive years for a total budget of $4M over the three year period.

Successful completion of the above work would result in an ab initio ability to implement a 100 million device logic circuit literally overnight, moreover under beginning-to-end control of a digital computer. A supercomputer designer with such a computer-building tool could thus iterate his design at least once per day, several times more frequently than the most fortunate designer can attain at present. Alternatively (and perhaps more practically), a half-dozen designers could each cycle their independent supercomputer designs through such a system twice every week.

Appendix G. Monolithic Implementation of Supercomputers

Introduction

Take as a contemporary supercomputer the S-1 Project's Mark IIA uniprocessor. Consider its planned repackaging-with-extensions, the S-1 Mark V implemented on a single silicon wafer. How is this repackaging to be accomplished? The question may be naturally resolved into two components: 1) How is it possible to consider putting down an entire system onto a single wafer, just from a standpoint of available area; and 2) From where will the needed implementation technology be obtained?

Background

Why isn't every supercomputer built on a single piece of silicon? The advantages of doing so are seemingly compelling: the hundreds of thousands to millions of interconnections between small- and medium-scale integrated circuit packages used in such systems through the present time are the major source of system cost and unreliability (it's been authoritatively estimated that over 99.9% of chip failures are due to their peripheries, and that less than 10% of system failures are due to failures within an IC package).

The bulk and cost of packages and mounting for the tens to hundreds of thousands of chips used up to now completely dominate all current supercomputer system volume and mass budgets. For example, the mass of functional silicon in the ton-scale S-1 Mark IIA is of the order of 25 grams, and its volume is about 10 cubic centimeters. The Mark IIA's areal heat dissipation averages about 15 watts per square centimeter of IC package, less that two percent of the levels which Tuckerman and Pease have demonstrated experimentally can be removed from silicon continuously and with low temperature drops.

There is apparently only one reason why anyone builds supercomputers out of more than a single piece of silicon: it is widely believed that monolithic implementations of supercomputers are currently not feasible.

From where does this belief originate? From the completely outdated notion that the required area of silicon is huge, which ceased to be valid when lithographic line widths began approaching one micrometer (μm) in scale, and from the belief that the density of implementation defects--currently about 3-5 per square centimeter for commercial NMOS processes--precludes practical implementation of any circuit requiring over about a square cm in area.

That the intelligent use of reduncancy can obviate the impact of defects has been widely understood (though little used) since von Neumann's classic studies of a third of a century ago; the associated 'redundancy insurance premium' imposed on digital logic is relatively small (e.g., slightly more than three-fold, for majority-voting logic approaches), while that for digital memories has repeatedly been demonstrated to be of the order of a few percent for redundant rows-and-columns approaches, and is of the order of the natural logarithm of the populations of strings-of-bits for Hamming-like techniques applied on a system scale. Disposing of the effects of defects is thus entirely within the purview of the digital system designer, whose present tools are more than sufficient to deal with them (at least in principle). The question of implementation feasibility is discussed in quantitative detail below.

Allocating the Wafer Area

The semiconductor surface area needed for monolithic implementation of a supercomputer-scale system, such as the S-1 Project's Mark IIA uniprocessor, can be resolved into seven basic portions: those for the CPU, the on-wafer high-speed memory not intrinsic to the CPU, the on-wafer operating system, the on-wafer I/O interfaces and controllers, the buffering of signals to and from the off-wafer world and the associated wire-bonding pads, and the diagnostics/maintenance processor.

Take the unit of length to be the minimum linewidth attainable with the lithographic process used for implementation, and the area unit to be the square of this length unit. Assume that, for system reliability reasons, all blocks of memory will have redundant rows-and columns for initial obviating of manufacturing defects, and that system-scale data storage will also have double-error-correcting-triple-error-detecting (DECTED) coding. Both of these redundant encoding features increase the total number of bits required to store a unit of information.

The CPU is taken to consist of one million 2-input gates-equivalent (both the S-1 Mark IIA and the Cray-1 use roughly a half-million equivalents of such gates). These gates typically require 100 area units to form and usefully interconnect, so that one hundred million area units will be required for the CPU's logic. Assume that the CPU logic will be diced into sufficiently fault-free modules during the design process for subsequent discretionary interconnection such that it is sufficient to replicate each and every such module four times to insure adequate fault-tolerance over the wafer system's full operation lifetime; this will suffice, for instance, to disconnect one faulted member of each type of module during post-creation checkout and to have the remaining three work in majority-voting logic thereafter. Thus, 400 million area units will be required to implement the fault-tolerant CPU.

However, most of the active devices in the CPU are used to constitute static RAMs for storage of microcode, CPU state, cached data, register functions, etc., in about that order of numerical dominance. The S-1 Mark IIA uses a total of 7.1 megabits of such RAMs, of which designer estimates of the minimum required fraction range from 0.4 to 0.7. Assume that the S-1 Mark V will feature larger stores of all types, but will use them more effectively, so that 10 million bits of static RAM capacity will suffice for its CPU-internal stores. A static RAM can be implemented in about 70 area units per bit, so that 900 million area units are needed for such storage functions in the S-1 Mark V (allowing 28% or 200 million area units, for row/column redundancy, periphery overhead circuitry, and DECTED-ing of the long words used in the CPU). The [logic+memory] total requirement for the comprehensively fault-tolerant CPU is thus 1300 million area units, and is dominated by memory requirements.

How much high speed memory external to the CPU is appropriate to go on wafer of course determines the area required to implement it. Take two million words as being a not completely gross violation of Amdahl's Rule (which would suggest about ten times as much--100 megabytes--as being appropriate for the 100 MIPS Mark V). Inmos workers describe a fast 64K dynamic NMOS RAM with a cell size of 18 resolution elements (only twice the theoretical minimum). Thus, with a 26% increment for block- and system-level redundancy and RAM block overhead (periphery circuitry such as address decoders, sense amplifiers, output drivers, etc.), 1000 million area units are needed per million DECTED-ed 36-bit words of dynamic MOS RAM implemented, and 2000 million area units are needed for the two million words of on-wafer RAM.

The on-wafer operating system and its associated software (language support, etc.) may reasonably expect to require about 4 megabytes of EAROM-type storage: Ada subset compilers currently on the market require between 100 and 256 Kbytes of storage, and available estimates indicate that the S-1 Amber uni-/multi-processor operating system will be a few megabytes in size. Motorola workers presented at ISSCC '82 a fast EAROM with a cell size of 28 area units. With a 33% excess for block- and system-level redundancy and periphery circuitry, four megabytes of EAROM storage will cost 1200 million area units.

The area requirements of the I/O Data Stores which serve to buffer the very high bandwidth Mark V channels have been budgeted for in the CPU section. If they are to be provided with on-board supervisory processors, these will require about 8 million area units each, if they are comparable to the Motorola MC68000 microprocessor in complexity (whoe popular 8 MHz form is implemented in 4 μm linewidths and occupies 0.33 square cm, exclusive of bonding pad area, or about

2.1 million area units), and are implemented in triply-redundant majority-voting logic. Eight such ports will require 100 million area units, including 50% extra for interconnections to the CPU and to the output areas on the wafer periphery. UARTs have an area cost of less than 100,000 units each. On-wafer 14-bit DACs (with sample-and-hold amplifiers) may be accommodated at the cost of 275,000 area units each; both unit quotes are for bipolar implementation processes, but include bonding pads on the peripheries. It thus seems clear that eight units of each of these types could be included on wafer for a total area cost of 10 million units.

Off-wafer I/O buffering and bonding pads are closely associated in function and geography, and may be considered together. Industry-standard bonding pad size is a 4 mil square, and a 1 μm channel length MOS transistor of this scale can source/sink 0.5 amp; while such a transistor will clearly require a train of drive transistors to exponentially down-shift the impedance of a signal coming into it to be driven off-wafer, it will also logarithmically dominate their aggregate area. Implementation of a pair of robust protective diodes between each such pad and its buffering transistor (train) will require comparable areas of 16 square mils each, so that of the order of 100 square mils will be required for each input/output pad and associated buffer circuitry: one 16 square mil allocation for the pad, for each of the two diodes, and for the buffering transistor, with 20 square mils of connecting and isolation room.

Use of 80% of the periphery of a 125 mm (5 inch) diameter wafer can accommodate 1200 bonding pads on 10 mil centers (with the remaining perimeter being reserved for power connections). This would accommodate, for instance, 20 full word-width buses (36 bits plus 4 bits of per-9-bit=byte parity), plus 30 byte-wide buses, plus 60 serial lines plus 100 interrupt lines; this seems ample for essentially all conceivable military computing applications. This, or any comparable, [pad+diodes+buffer] area allocation will require 120,000 square mils, or 0.80 square cm of wafer area.

The diagnostics and maintenance processor will be taken to be comparable to a MC68000 in complexity (though implemented with majority-voting logic throughout). However, it will have far larger memory requirements: it must remember what the correct processor state is for many different test-generated circumstances. It will be assumed that 100 complete processor states (or their equivalents in a larger number of sub-states) must be recalled for diagnostic purposes, and that 100,000 bits characterize each state; thus 10 million bits must be recalled, moreover from EAROM. (The number of multiplexer/demultiplexer values that must be set within all the other major systems on the wafer at each power-up time by the diagnostics/maintenance processor by recalling previous test results from its memory in order to select healthy modules or de-select faulted ones is likely to be far less than a million.) From the discussion above regarding operating system storage on-wafer, it is clear that approximately 400 million area units will be required by the diagnostics/maintenance processor for its EAROM storage, which dwarfs the area requirement for the processor itself.

A summary of this wafer area budget is as follows:

Subsystem	Area, megaunits
Central Processing Unit, S-1 Mark V	1300
On-Wafer Main Memory, 2 million words	2000
On-Wafer Software (Amber, etc.), 4 megabytes	1200
I/O Controllers, w/one DAC and ADC unit each, 8 ea.	110
I/O Buffering and Connection Pads, 1200 lines; 0.8 cm^2	*
Diagnostics/Maintenance Processor	400
Total	5010

*Area unit-dependent; small in any case

The Implementation Technology

Recall that in February 1982, Intel workers introduced at ISSCC their HMOS III technology, which features minimum linewidths of 1.5 µm, and in which commercially available components are presently offered, e.g., their 20 nsec access time CMOS-centered 4K ECL RAM. Also at ISSCC '82, Martin Lepselter of BTL described a 0.5 µm x-ray linewidth lithographic process for NMOS with a demostrated 0.1 µm effective channel length which has already been used to produce two VLSI chips, a 40,000 device e-beam machine controller and a 50,000 device real-time speech recognition system. This BTL technology has already been put into use by Western Electric for mass production of 0.2-0.3 µm effective channel length VLSI devices. It is interesting to note that the BTL technology builds wafer-scale masks and exposes a resist-coated wafer monolithically (moreover, with a 15 resist exposure period) in a system whose reproduction cost is $300 K, and which is available within the U.S. on a no-fee cross-licensing basis.

Intel offers their HMOS III technology through their silicon foundary service, though Bell/Western Electric will be employing its x-ray lithography-based technology for proprietary production only. However, at least one Silicon Valley company is offering a full sub-micron x-ray process. Wafer-scale integration is already a reality; Burroughs has created a self-testing and organizing 'smart memory' with such technology.

Thus, if the 1.5 µm linewidth/2.25 square µm minimum area, commercially available HMOS III technology were to be used for its implementation, the 5010 million area unit Mark V computing system scoped out above would require 112.7 square cm of wafer area; in a 0.5 µm linewidth/0.75 square µm x-ray lithographic technology, building this system would require 12.52 square cm of wafer surface. For comparison, a 125 mm (5 inch) diameter wafer has an area of 122.7 sq. cm, while a 75 mm (3 inch) diameter wafer has a 44.2 sq cm area. An augmented Mark V system with a full Amdahl's Rule of on-wafer high-speed memory (128 megabytes) and twice as much (8 megabytes) on wafer software storage would occupy 90 sq cm of wafer area if implemented in 0.5 micron linewidth technology.

Power and Cooling Requirements

What does it cost in power and energy to operate such a system? The Intel HMOS III technology features a speed-power product for its gates of 250 femtojoules, while that of the BTL short-channel gates is about 40 femtojoules; the Intel gates have minimum delays of 200 psec, while the BTL ones are at the 30 psec level. Thus, the BTL gates draw about 1 mW of power, while the Intel ones require about 1.25 mW. If implemented in CMOS technology, rather than NMOS, roughly one-third to one-half as much power would be required to operate typical CPU logic at the same nearly limiting speed, due to the necessity to switch only a fraction of the gates in any cycle.

Thus a million gate CPU would require of the order of a kilowatt of power supply, if implemented in either the Intel or the BTL technologies; of course, the BTL one could be expected to exhibit a cycle time of the order of 1 nsec (e.g. gate delays, or a scalar computing rate approaching 1000 Million Instructions Per Second (MIPS), while the cycle period of the Intel one would be of the order of 6 nsec, corresponding to a scalar execution rate in excess of 100 MIPS.

If CMOS variants of these technologies were to be used, 300-500 watts of power would suffice for the CPU when operated at full speed, and proportionally less at lower clock rates. The 10 megabits of high speed static RAM carried within the CPU will dissipate a fraction of this power. The contemporary Intel CMOS-internal/ECL-external static RAM has an equivalent address access time (deducting the ECL conversion times) of about 10 nsec with 2 μm channel lengths (3 μm lithography) and dissipates less than 50 microwatts per bit stored, when cycling at full speed. In 1.5 μm technology, the corresponding 10 microwatts/bit would entail a dissipation due to all of the CPU-internal storage of about 100 watts.

One might choose to implement at least the EAROM for software (and diagnostics/maintenance processor data) storage in CMOS RAM, in which case the 4 megabytes of DECTED storage would require only 5 microwatts of power to retain data, if implemented in the new Hitachi 0.1 picowatts/bit CMOS RAM technology (op. cit., p. 260). (If the 10.5 square inch area, 0.07 inch thick Panasonic lithium/polycarbon/monofluoride battery were to be used to power this system in data retention mode, its 1.5 amp-hour capacity would operate it for 300,000 hours, about ten times the 30-year shelf life of this battery technology. This same power supply would operate it at 10 MHz accessing rates for several minutes. However, since this store contains software which would presumably be accessed infrequently, such rates would likely not be sustained for such periods; the 6 watt peak dissipation of this store would be sustained for very brief periods, as the process refreshed its caches.

Contemporary dynamic RAMs implemented with 2 μm linewidths require about one microwatt per bit when being cycled, and about 50 nanowatts per bit during standby operation. Pure NMOS scaling suggests that these figures should decrease as the square of the linewidth (but is of doubtful applicability over a very large range). Due to the presence of the caches within the Mark V processor, the main memory should be cycled very infrequently--less than one percent of the processor cycles--so that refresh requirements should dominate the DRAM storage current requirements. The on-wafer high-speed memory capacity of 100 megabits should thus dissipate 5 watts, if not implemented in any better technology that that already being sampled commercially by Hitachi; recent technological trends indicate that at least an order-of-magnitude improvement in this per-bit dissipation should be available by the time the S-1 Mark V is implemented. If the 100 megabits of cache-backing store were to be implemented in CMOS RAM, rather than DRAM, the standby operating power requirement of 10 microwatts would be negligible by any standard; the 6 watts of peak power requirement at 10 MHz operation would be sustained at most 1% of the time, yielding an average power draw of the entire on-wafer memory of the order of 0.06 watts.

Dissipated power has been removed from silicon wafer backs at flux levels greater that 700 watts/square cm via forced convection with junction to water temperature drops of less than 30 kelvins; more recent variants of this basic technology appear ideal for high density heat removal from any type of silicon wafer. A planar version of a water-loaded heat pipe can remove heat from a silicon surface in good contact with it at a rate in excess of 150 watts/square centimeter, and advect it through heat dumping channels at heat-equivalent flux levels of several kilowatts/square cm. The maximum of the heat dissipation values just estimated is in the CPU portion of the S-1 Mark V wafer, and is of the order of 200 watts/square cm in 0.5 μm technology, and 20 watts/square cm in the 1.5 μm lithographic resolution technology. Both of these values are substantially less than that processed by the demonstrated Tuckerman-Pease technology, and are at most comparable to those processed by low-technology, and are at most comparable to those processed by low-technology heat pipes.

It therefore seems straightforward to cool the most intensively heated portion of the S-1 Mark V wafer, using existing technology. Required power input is about a kilowatt for CMOS technology. Projected performance levels for the CMOS version are 5 to 30 times that of a Cray-1 or Mark IIA (depending significantly on the lithographic resolution employed in silicon patterning). A minimum specific performance level of 10 Cray-1s per kilowatt inputted may therefore be expected.

Operation in the Military Environment

Three features not found in the civilian environment of high performance computer systems are of great importance in the military one: tolerance to high levels of penetrating ionizing and particulate radiation, to intense electromagnetic fields and to high accelerations. Monolithically implemented supercomputers, such as the S-1 Mark V, can be made more tolerant of such environmental stresses than are current military electronics of all types, simply because of the greater feasibility of completely enclosing them in strong, electrically conducting, high opacity shells. Doing so is far less feasible for current military computer systems, because of their large size and relatively high power dissipations. These problems will be obviated with future monolithic supercomputers which will be less than a centimeter thick, a dozen centimeters in diameter and dissipate fractions of a kilowatt.

Conclusion

Nothing presently stands between digital systems creators and monolithically implemented supercomputers. All that is required is complying with the current technological imperative with the means already at hand, and the resources to do so.

VII. BIBLIOGRAPHY

There are a number of papers and reports that have been stimulants in developing the programs and plans described in this paper. Those referred to in the text are listed below.

A. Wood, L., "Concerning Technical Means for Dealing with Aspects of the Near-term Information Onslaught," UCRL-86581, June, 1981.

B. Wood, L., "Synthetic Aperture Radar Data Processing: The Challenge and Opportunity Posed by Orbiting Systems," UCID-19747, September 1982.

C. Wood, L., "Conceptual Basis for Defense Applications of Supercomputers," UCID-19746, December 1982.

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government thereof, and shall not be used for advertising or product endorsement purposes.

CONCERNING TECHNICAL MEANS FOR DEALING WITH ASPECTS OF THE NEAR-TERM INFORMATION ONSLAUGHT*

Lowell Wood
University of California Lawrence Livermore National Laboratory
Livermore, California 94550

I bring fraternal greetings from the staff of the Lawrence Livermore National Laboratory where, as in all other Laboratories of the University of California, untrammeled intellectual inquiry thrives, except on certain rare and regrettable occasions.

---

*A paper prepared for presentation at *The Conference on Science and the Information Onslaught*, Los Alamos National Laboratory, Los Alamos, New Mexico, 1-5 June 1981. Work performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under contract number W-7405-ENG-48.

From Where Is 'The Information Onslaughting?

When considering near-term means for coping with the information onslaught, a prudent student will inquire a bit beforehand as to its nature; when commencing to do so, one is drawn naturally to examine the dominant features of the beast. Guided as always by Lord Kelvin's admonition that the distinction between philosophy and science is that the latter intellectual endeavor has numbers associated with it, one starts inquiring for the quantitative aspects of "the charge of the bits."

If one judges by size scales alone, the information onslaught is perhaps to be identified most aptly with the arbitrarily large amounts of information available from integrating the differential equations describing physical reality to generate histories of particular physical situations characterized by their initial and boundary conditions. It is only little realized even today that the advent of digital computing technology has elevated physics in less than a single human lifetime from natural philosophy to an ability to predict the future of patches of physical reality with degrees of fidelity, detail and omniscence completely beyond the comprehension of the Delphic oracle. To the most profoundly humbling question posed to Job three millenia ago, the human race can now answer 'Yes!', due in large part to the elaboration of physical theory which modern digital computing has made possible. The huge intellectual advances of the past few centuries have brought us knowledge of 'the ordinances of heaven,' and digital computing has taught us much of 'the dominion thereof in the earth' (Job, Chapter 38).

Corresponding to this capability and the value which at least some national governments have attached to it, many of the largest computer systems in the world have been devoted to such scientific soothsaying for the past third of a century. Indeed, it is worth recalling that electronic digital computers were born during the largest scale human conflict in history out of the hope that their possession would confer major advantage, if not victory, and that their youth was generously nurtured for similar reasons when the conflict shifted to less violent forms. At the very least, such a recollection may be well-received at this Laboratory, where modern digital computing undeniably came into its own under the leadership of pioneers such as Nick Metropolis and his colleagues.

The suggestion that physical modelling activity could be at the center of the information explosion gains plausibility when it is considered that there are hundreds of thousands of practicing senior scientists and technologists alive, many of whom are potentially curious as to how a physical situation of interest to him or her will behave under varying circumstances, but who lacks at least some of the resources required to create and study the real situation as desired. Each is thus a potential customer for a super-scale computer system for digital modelling purposes, probably one comparable to the most powerful currently in existence. Moreover, as modelling completeness and integration accuracy can only be finite on any finite speed computing system, human proclivities toward perfection may guarantee an open-ended demand for such information processing and manipulating capability.

As a specific example, the scientists at my Laboratory at Livermore were provided with digital computing capacity which doubled in effective capacity an average of every 16 months over a two-decade period during the Laboratory's youth, and found this situation very productive; indeed, the cessation of this exponential growth in capability during the past decade was quite traumatic. Moreover, the Livermore experience was not atypical, either in its exponential growth curve during the '50s and '60s, or its relative stagnation during much of the '70s. Sustained exponentiation over many decades certainly is one of the central characteristics of an explosion — perhaps this explosion in information processing capacity for physical modelling purposes is the center of the information onslaught. It's as close as I have been able to locate, and the rest of my talk will take it as a point-of-departure.

All Dressed Up And No Idea Of Where To Go

The currently most powerful computer system for physical modelling purposes is the CRAY-1, which has an incremental production cost of about $3 M. (I remind you that a CRAY-1 does high precision arithmetic at peak sustainable rates comparable to that which could be attained by the entire human race together, each using one of the best commercially available hand calculators; on the other hand, performing high precision arithmetic, even with powerful aids, is certainly not what the human brain does best.) A moderately fortunate and serious physical modelling researcher presently gets use of about 3 percent of a CRAY-1 on the time average; that is, access to digital modelling tools is presently allocated so that prorated incremental production costs of modelling hardware are comparable to the modeller's annual burdened manpower cost to his organization, which is not retrospectively surprising. At the Livermore Lab, those of us in the S-1 Project are presently building a set of computers, each one of which is comparable in physical modelling power to a CRAY-1 but whose incremental production costs are an order-of-magnitude lower, due to the last half-decade's technological advances. This implies that an order-of-magnitude increase in computing power will eventually become available to the average researcher, which will serve to whet appetites — if not completely satisfy them — for 3-dimensional physical modelling at space-and-time resolutions comparable to present 2-dimensional work.

We expect the S-1 work to culminate in the mid-80s with the realization of a computer system with an order-of-magnitude greater computing power than that of a CRAY-1, realized on a single crystal of silicon less than a dozen centimeters in diameter. It is interesting to note that *all* of the technology necessary to realize such a system is currently available, with just one exception: we do not know exactly how to employ the exceedingly powerful tools which are presently available for patterning silicon.

A few technological considerations may serve to illustrate this tantalizing problem. Present state-of-the-art computer systems contain anywhere from several million to a few dozen million transistors in their central processing units, corresponding to several hundred thousand logic gates or switches and as much as several million bits of Central Processing Unit internal 'state' or specification-of-processing-condition. Each of these gates typically switch information after an intrinsic delay of most of a nanosecond, permitting the CPUs which they constitute to perform an operation in one to two dozen nanoseconds — the characteristic couple dozen gate delays per major computer operation. Now virtually everyone has heard of quantum interferometric ('Josephson junction') digital logic gates operated at liquid helium temperatures which can switch on 25 picosecond time scales; less widely hearalded are room temperature silicon gates which have been demonstrated to operate with 40 picosecond gate delays, moreover implemented in *the* standard commercial mass-production VLSI technology, n-channel MOS. To completely process a full 10 centimeter diameter wafer in this latter type of integrated transistor technology involves a present cost of about $50; a wafer of this size could carry a few of the mid-80s S-1 computer systems to which I have referred. The implications of integrated circuit physical implementation capabilities are utterly manifest in the Figure: continuing of the past decade's merely exponential advance in time of circuit density (and speed, and cost-effectiveness, and energy efficiency, and ...) is the *least* that can be expected during the '80s. It is therefore difficult to see from where the production cost of a state-of-the-art superprocessor will come, a half-decade hence; the cost a decade from now is completely unclear from the present vantage point, a remarkable state of affairs characteristic of revolutionary circumstances.

But why the half-decade delay? If we can even now paint frescos on silicon which not merely delight the eye but emulate major aspects of the brain, what is holding up the advance of this most modern — and most significant — art?

The basic problem is that we have moved all too rapidly from an era of silicon famine — that is, the period in technological history characterized by acute, cost-driven scarcity of logic and memory in computers — to the time of silicon plenty, in which the large-scale integration revolution has already driven the incremental mass production cost of the highest speed transistors below $\$10^{-4}$ each. While we have long known in general what to do when finally entering this promised land, in characteristically human fashion we really did not expect it to happen quite so soon, and we certainly did not adequately prepare for it.

Birth Defects: The Chock in Front of the Wheel of Progress

In retrospect, the genesis of this problem occurred when the advent of the earliest computers successfully drove the improvement of vacuum tubes to performance and reliability levels sufficiently high that computers containing thousands of them could operate for a few hours between comprehensive maintenance periods. This unfortunate — and quite unexpected — technological advance allowed the early computer designers to circumvent the otherwise compelling necessity analyzed by von Neumann and his collaborators to create reliable computing systems out of unreliable components, through the intelligent use of redundancy. Thus, in order to save something of the order of a factor of two in computer construction costs during the past third of a century, we have foregone having highly reliable, single-point failure-immune computer systems during this same period. Though clearly tolerable for this entire period, single-point failures are currently becoming *the* impediment to further technological advance. Now that the incremental construction cost of a state-of-the-art computer is approaching the cost of a single man-day of maintenance technician time, the economic imperative has flipped direction and presently points the same way as the esthetic one always has: towards single-point failure-immune computers.

There is an extremely important practical consequence of single-point failure-immunity of an operating computer associated with the extension of this feature all the way back to when the brand-new computer is first turned on. Recall that most modern microprocessors — the ancestors of supercomputers-on-a-chip — are born dead; due to the presence of one or more fabrication flaws of an irreparable nature, they are inoperable in some significant sense. This is well-known to semiconductor specialists as "the yield problem". Through the present time, the power of microprocessors is limited by how much circuitry can be integrated together before the likelihood of a fatal fabrication flaw appearing at random within its implementation becomes too close to unity, i.e., the 'yield' of fully functional chips from a wafer of them becomes too low. However, a single-point failure-immune supercomputer-on-a-chip may have such congenital defects, but they will be masked — made effectively invisible from the outside of the chip — by single-point failure-immune design of the chip's circuitry.

Indeed, this immunity to a single birth defect is already more the rule than the exception in the most modern semiconductor memory chips (64 K static RAMS). Containing up to a few hundred thousand high speed transistors per chip, typically 97 percent of these chips are stillborn, with at least one fatal manufacturing flaw somewhere on them. With the built-in, highly automatically implemented ability to substitute a new set for a single memory cell's half-dozen transistors, 80 percent of these dead-at-birth chips may be revived to lead useful lives of full duration and capability. Interestingly, the increment to the transistor population of a memory chip to raise live chip yields from 3 to 80 percent is typically less than 2 percent — the intelligently applied redundancy tax is really very low!

The huge gain in cost-effectiveness associated with such minimal redundancy is indeed striking; on the other hand, it is intrinsically limited to computer memories, and CPU logic may have to be duplicated or triplicated in order to become comparably robust against all first contingencies. That is the bad news — the good news is that over 90 percent of the transistors in a modern CPU are actually used to constitute memory elements, with only ten percent or less devoted to making logic gates: modern computing structures are really evolving rapidly towards intelligent memories, with large bricks of memory being mortared into computing structures with small dabs of logic. The pervasive use of redundancy to safeguard against the deleterious effects of all first — and quite possibly even multiple — contingencies in computers to be created during the next half-decade is thus likely to increase their total transistor populations by a few tens of percent, and certainly by less than two-fold. Since the number of transistors which can be created on a given area of silicon doubles in little more than a year these days, the 'premium' in transistor population required to realize truly failure-immune supercomputers-on-a-chip will be readily affordable, particularly since the alternative is manufacturing yields of a tiny fraction of one percent and frequent, irreparable failures-in-operation of the miniscule fraction which are born functional.

The work of the next few years in realizing supercomputers-on-a-chip will thus be focussed primarily on creating highly automated means of introducing the required degrees of redundancy into supercomputer designs, and secondarily on projecting the resulting designs onto monolithic, nearly two-dimensional semiconductor-based surfaces. The S-1 Project, as well as similar undertakings elsewhere, sees these two tasks as *the* challenge to computer technology in the early '80s. As I remarked before, realizing such designs in extremely high performance silicon can be accomplished with technological means already in hand, e.g., discretionary e-beam direct-write-on-wafer systems generating fractional micron effective channel length nMOS transistor patterns, interconnected in the usual fashions.

Appropriate Behavior Upon Entering the Promised Land: MIMD Networking

So here we are a half-decade hence with the ability to crank supercomputers-on-a-chip out of our silicon foundries at least as agilely as a short-order cook whips up a stack of pancakes. Every serious professional has ten CRAY-1-equivalents of personal computing power in his or her office later in the decade, at a cost comparable to that of current high-performance hand calculators. Is this the extent of the potential of the ULSI revolution, the full implication of attaining the Ultra-Large Scale Integration plateau? Is there no practical way to increase effective computing power faster than the underlying logic technology improves?

The S-1 Project, along with a growing band of others elsewhere — and I was delighted to hear yesterday that Jack Schwartz is among them — believe that confederations — federal unions, if you will — of N nominally independent processors can be designed, built and programmed to support reasonably close to N-fold speed-ups in the wall clock rate at which most problems of interest can be solved.

Very straightforward considerations of time-average computing efficiency of a typical processor suggest — as Jack sketched yesterday — that one should have as much computing power in each single member of the multiprocessor federation as can be provided in a reasonably cost-effective fashion, and so the first S-1 multiprocessor is being implemented during the coming year with processors of roughly the scientific computing rate of a CRAY-1.

Considerations of proper balancing of hardware and software costs suggest that an experimental multiprocessor system should be of sufficiently great performance scale to justify the costs of developing serious systems and applications software for it, and so the first S-1 multiprocessor is being constituted of 16 processors.

Similar considerations suggest the appropriateness of as high a quality interconection between processors and memory moules as is economically feasible, and so the highest connectivity switching network — a full crossbar switch — is being provided to simultaneously interconnect all processors in any one-to-one-onto mapping to all memory modules in the S-1 multiprocessor, to arbitrate fairly in the event of momentary request conflicts, and to provide similar interconnectivity of processors among themselves.

It is worth noting in this latter respect that full crossbar switches are an eminently practical means for connecting together even quite large federations of processors and memories. While those who study data switching networks professionally find crossbars too simple to provide much grist for their mills, and thus disdain them in favor of more exotic interconections, technology quite perversely makes crossbars too cheap to forsake for more cleverly designed but intrinsically lower performance interconnection approaches.

As a specific example, we find that full crossbar connectivity of our S-1 16 member multiprocessor has a quadratically growing fraction whose cost is somewhat less than one percent of that of the multiprocessor system as a whole; in other words, full crossbar connectivity could be provided with our current S-1 technology to interconnect 10^3 processors with each other and with their memory modules before the cost of the switching network would rise to equal that of the processors and memory modules. Only when providing for networks of more than of order 10^3 processors with 1981 technology would one need to consider less costly approaches, such as the Perfect Shuffle idea which Jack sketched yesterday. Alternatively, one might choose to interconnect a number of networks, each constituted of several dozens to a few hundred crossbar-connected processors, into super-scale networks, again using crossbars to net the networks together.

But granted that one can build and interconnect such huge piles of hardware. Can such interconnections possibly be programmed with finite effort, can they be exercised with reasonable efficiency, and can their enormous number of possible failure points be coped with? Interestingly enough, the answer appears to be 'Yes' to all three of these questions.

Programming for the S-1 type of Multiple-Instruction-stream/Multiple-Data-stream (MIMD) multiprocessor turns out to be quite straightforward: a small number of grammatical constructs of a comment nature appended to the standard grammar of any high-level algorithmic language — such as FORTRAN or Pascal — suffices to extend it to an efficient multiprocessor programming language. Programs in this language may be readily compiled for execution on a multiprocessor with a processor population which is dynamically variable from 1 to N *during* program execution. The increase in programming effort to fomulate problems for a multiprocessor, relative to that for a single processor, is still being investigated, but appears to be of the order of 10 - 25 percent for typical scientific algorithm mixes.

Multiprocessor utilization efficiency studies also turn up preliminary but quite optimistic results. Examinations of both analytic and simulation natures indicate that a reasonably wide variety of scientific algorithms can be executed on a 16 member MIMD multiprocessor of the S-1 variety with average hardware utilization efficiencies of between 55 and 95 percent. The overall time-averaged efficiency of MIMD multiprocessor execution of a Livermore Lab benchmark two-dimensional Lagrangian shock hydrodynamics and ADI heat transport physical simulation code was found to be 75 percent, i.e., it ran 12 times faster on a simulated 16 member S-1 multiprocessor as it would have executed on a single member processor. Jack quoted closely comparable results to you yesterday, and still other researchers find similarly. Clearly this is not a universally applicable number, but it is representative of mainstream scientific computing, and as such is very hopeful.

The Gordian knot of the multiprocessor reliability question can always be sliced apart by checkpointing a calculation to redundantly reliable memory adequately often, and then exploiting the variable processor population feature of the multiprocessor's problem partitioning software to always use all of the multiprocessor's healthy processors. More elegant approaches — akin to untying the Gordian knot — which are too complicated to sketch in this short talk also appear to be practical.

In summary, the MIMD type of multiprocessor appears to open up a track of a highly practical nature leading to comparatively huge increases in effective computing power, even though the performance of the underlying logic may improve only relatively slowly. As such, it is *the* method of choice into the foreseeable future for translating the continuing enormous gains in digital logic density and cost-effectiveness into correspondingly great advances in useful computing power. I therefore suggest that it will be at the center of means of dealing with near-term aspects of the information onslaught.

Meanwhile, Back on the Hardware Frontier ...

However, likely advances in the underlying hardware technologies should not be ignored. Digital logic gate speeds seem likely to advance by at least an order of magnitude from the 25-40 psec rates at which the fastest JJ, GaAs and Si gates work at present, though exploitation of the one psec JJ switching rates very recently reported, by the development of 'nearly short' channel GaAs MESFETs comparable in scale to the currently most advanced Si gates, and via continuing advances in fractional micron Si lithography. Advances to sub-picosecond gate delays via optical digital logic, though of higher technical risk, nonetheless seem likely in the '80s through work underway at Livermore and elsewhere.

It is also crucial to note that the energy needed to perform an elementary logic operation, presently a few femtojoules, continues to decrease rapidly with technological advance. There is still a gap of a million-fold between the energy dissipation of the best digital logic gate technology and that theoretically required. Closing most of this six order-of-magnitude gap with molecular-scale transistor elements which are feasible to fabricate and interconnect in exceedingly large quantities is a surpassingly exciting prospect, one being pursued at Livermore and probably elsewhere.

It is easy to forget that memory is part-and-parcel of computing, and that a powerful CPU with insufficient memory attached to it is like a weight-lifter with paralyzed legs: a few feats can be performed impressively anyway, but in most performance is degraded, and some things are just impossible. Indeed, one of the cardinal empirical relations of modern digital computing, Amdahl's Rule, states that a computer must have approximately one word of high speed memory attached to it for every instruction per second of computing which it is expected to efficiently perform. It is just because memory is so regular, so symmetric in nature, that it is easy to design and efficient to manufacture; memory technology has characteristically paced digital semiconductor technology development as a whole. We thus tend to take it for granted, relative to the "problem child," digital logic, with its highly irregular, quasi-random nature.

High-speed semiconductor memory is presently quite fast — data recall times range from somewhat under 10 to about 100 nanoseconds — and very cheap — about $\$10^{-2} — 10^{-4}$/bit (depending on the required speed of data recall). Moreover, another famous empirical rule due to Noyce, the co-inventor of the integrated circuit, states that memory price and the silicon area required per bit stored drops by two-fold annually, on the average; this rule has been in effect for the past decade, and continues to be valid through the present.

Interestingly enough, some types of semiconductor memories require no power to retain data (but can only be written perhaps 10^5 times, though read as often as desired), others retain data without power but can be re-written only quite slowly (though read very rapidly and arbitrarily often), while still others are completely high speed read/write units and only require a few nanowatts per bit to retain data indefinitely. These properties are *all* of interest in creating appropriate memory systems for very high performance computer systems for a reason both rather profound and pervasive: computers, like the people who program them, tend to process data in a peculiarly orderly fashion. Data being processed at any particular moment tends to be closely associated — by some memory metric or another — with data just processed and with that about to be processed. Perhaps such time-sequenced processing of mostly 'nearby' information is intrinsic to most all intelligent manipulation of data; somewhat more likely, it's a subtle feature of how human brains store or retrieve data.

In any event, computing almost exclusively on 'nearby' data is the way it's observed to take place in nearly all applications, and it permits the extraction of 'locality' from data masses and the exploitation of such locality via hierarchies of memory having differing costs, power requirements and recall time features. Each lower layer of such memory hierarchies — starting from the top one serving immediate CPU needs — has exponentially greater storage capacity and greater required time-to-recall than the layer just above it. The effect of a very capacious and inexpensive memory with only minimal penalties in *average* recall time for a datum are thereby gained: typically 99 percent of all data needed by the CPU are found in the fastest memory, and typically a 10-fold penalty in retrieval time for one percent of the data requests (or a 1000-fold penalty for 0.01 percent of the data requests) imposes only a 10 percent average latency penalty while buying increased memory capacity at a 100-fold lower per-bit cost, in moving from one layer in the hierarchy to the next lower one. Very many types of memory technology thus find natural niches in this hierarchy, and enrich its potentialities in the process.

But how much memory capacity of what speed and unpowered data retention characteristics will the computer systems of the '80s need, and how much is technologically accessible? Amdahl's Rule provides a time-tested lower bound, and John McCarthy has suggested a reasonable upper bound: a computer should be able to recall its entire previous state from the time it was first put into operation, akin to the most legendary human recall capabilities. Since a 1981 state-of-the-art computer changes an instantaneous state size of the order of 10^2 bits about 10^8 times per second over a really useful life of the order of 10^8 seconds, this would require a memory capacity of the order of 10^{18} bits.

I will state without proof for reasons of brevity that it is possible to construct with existing technology a photochemically based, molecularly encoded memory system having an order-of-magnitude larger capacity — an Einstein of bits — which could be retrieved on sub-microsecond time scales and whose active medium would be a fractional cubic meter in volume. If you are more conservative and inclined to just write a purchase order for a more modest system, a very large and most reputable high technology company is prepared to deliver a functional prototype of a 10^{15} bit capacity, photochemically-based storage system for something of the order of $30 M by mid-decade with potentially a microsecond recall time, as at least some of you are already aware. By way of comparison, I will remind you that a human who could perfectly recall absolutely everything which he or she sensed during a typical lifetime would have stored something of the order of 10^{14} bits; most of us recall something of the order of 10^{7} bits on 1-10 second time scales with high error rates, and, even more raggedly, perhaps 10^{10} bits total, over 10-10^{5} second recall periods.

I would say more on this subject, but Edward will be treating it exhaustively in his talk on Thursday evening, and prudence — if not courtesy — requires that I forebear.

Now That You Have Everything You Ever Wanted . . .

So we can realize at mid-decade computing systems on a single chip which are an order-of-magnitude more powerful than the best ones at present for an incremental cost not greatly in excess of $10. So we can interconnect many hundreds of them intimately and for larger numbers somewhat less closely. So what? Everyone then has 100 million Floating Point Operations Per Second — 100 megaFLOPS — of computing power in his/her office terminal, whose cost, mass and power consumption are probably dominated by the visual display unit. Is this where the leading edge of the information onslaught will be in the second half of this decade?

I suggest that, in terms of sheer megaFLOPS and Millions of Instructions Per Second, the answer is almost certainly 'Yes,' just as microprocessors, rather than CRAY-1s, represent the bulk of planetary computing capability at present. However, the seeds of the future will be sprouting elsewhere, with one such seedbed being the use of advanced networking techniques — in both hardware and software — to effectively organize large numbers of superprocessors-on-a-chip into single computing systems running single problems of sizes which are titanic, even by present scales.

A quarter century of observing and dabbling in the American computing scene is the basis for my remark that it is easy to get the Federal Government to put up $$10^{6}$ for a single computing system, challenging to raise $$10^{7}$ for such purposes, and effectively impossible to get Uncle Sam to spring for $\$10^8$ for any computer system; no other patron is in the game, as a practical matter. Incidentally, this little lemma has remained valid while the dollar has changed in real value by a half-order-of-magnitude, and has a corollary due to Seymour Cray, usually quoted as "A top-of-the-line computer system has always cost $10 million, and always will." When the day arrives several years hence that a CRAY-1-equivalent of computing power costs of the order of $1, what can be done with the million CRAY-1 equivalents of computing power that a relatively easy-to-get megabuck will buy?

I know what *I* will do with it, and I will sketch this briefly because I suspect that my tastes in physical modelling are not totally atypical. With two of the six orders of magnitude of computing power gained over the best present systems, I will buy adequate spatial resolution in the presently lacking third dimension. With two additional orders of magnitude, I will purchase a factor-of-three average improvement in space and time resolution in these four dimensions, relative to that enjoyed on present 2-D problems. With the fifth order-of-magnitude, I will get my simulation problems run to completion in a few tens of minutes rather than the several hours which they tend to require on a CRAY-1. With the final order-of-magnitude, I will buy a corresponding increase in the richness of the physical phenomena which can be modelled at a given logical point mesh over a time integration interval, thereby permitting *really* quite extensive population kinetics, or Schrodinger equation solving, or semiconductor device simulation, or whatever, to be performed. Note that such an allocation of aggregate processing power is compatable with having 10^5 interconnected processors at work — by hypothesis, there is something of the order of 3×10^7 logical mesh points in problems of interest, so that each of the 10^5 processors computes for an efficiently large number of grid points.

Others will choose to allocate the capabilities of such a $10^7/10^8$ megaFLOP system (depending on whether it is doing scalar or vector processing) quite differently, and it is clear that their freedom to efficiently do so won't be abridged by the MIMD multiprocessor structure, at least within very wide limits.

But what then? How will additional gains in multiprocessor hardware and software organization be put to reasonable use, to say nothing of the two additional orders of magnitude gain due to improvements in logic speed which can be foreseen during the '80s with fair confidence? I at least do not have particularly inspired suggestions. How to use the upcoming three-to-six orders of magnitude in effective computing power seems clear. How to reasonably employ a billion- to a trillion-fold more computing power — or memory capacity — than is currently available exceeds my imagination at present. Ask me again in a decade, please.

But Is There Life After 1991?

I would now like to briefly invite your attention to some of the implications of near-term advances in human ability to deal with the information onslaught. In particular, consider the magnitudes of the switching rates and memory capacities which we have just been considering relative to those of the human brain. The brain — and specifically its cerebrum — contains something of the order of 10^{12} neurons, each of which can switch at most about ten times per second; each of us can thus perform something of the order of 10^{13} elementary switching operations per second, a slightly awesome data processing rate, even for noisy logic elements. Assume now that each neuron makes switching decisions based on 10^2 bits of internal state — specifications as to how much rest it got the night before, the local glucose concentration and oxygen tension, how recently it fired, the present and recent state of excitation of its dendritic tree, how much and what particular flavor of memory protein, RNA or whatever it contains, etc. — recent research seemingly suggests that a neuron's state may be represented by more like 10 than 100 bits, but let's optimistically assume a relatively rich internal state. Note in passing that the ratio of memory bits to logic gates in the brain is the same order-of-magnitude as in modern computers CPUs, which I characterized earlier as ever more like intelligent memories — the bit-to-gate ratios of both information processing systems are of the order of a few dozen.

We have been considering superprocessors-on-a-chip whose 1981-technology gates (and associated interconnections) are about a factor of ten smaller in linear dimension than neurons, and which switch at least 10^9 times faster while consuming about 10^7 times as much power. A $\$10^6$ cost, late-1980s MIMD multiprocessor composed of such elements would contain 10^5 monolithically integrated superprocessors, each having 10^6 such switches and 10^7 bits of internal state; this is about 10 percent of the upper-bound estimates which I just cited of the logic switch and internal state bit counts of the human brain. If it is possible to allocate the decision-making work of 10 neurons to an average of a single silicon gate — just as we presently assign many logical mesh points to a single CPU in MIMD processing of physical models — and if 10^7 additional, non-state-specifying bits of memory are appended to each CPU chip, then this MIMD multiprocessor system might reasonably be expected to compute at least 10^6 times faster than a human brain: its ten-fold fewer switches each operate 10^9 times faster than do a neuron, and 100 gate operations conditioned by 3000 bits of total state is surely a *very* generous estimate of the complexity of a tenth-second of data processing operation of a single cerebral cortex neuron. Moreover, the (non-state) memory capacity of the total system would be adequate to store *all* the sensory information gathered by a human over an entire lifetime. Built in 1981 technology, this MIMD multiprocessor system would occupy about 10 cubic feet (including provisions for cooling), and would require about a megawatt of electric power. As the trend line in the Figure suggests, at least a thousand-fold reduction in these numbers by 1991 is to be expected.

I really cannot suggest to you what a million-fold increase in the computing rate of a human brain-equivalent would portend; I have only a hazy intuition as to what a ten-fold increase in human information processing rate would imply. Indeed, I have only been able to find one quantitative hint in this respect in the published literature, which rather cryptically suggests that God perceives duration — e.g, thinks — only a few hundred thousand times faster than do people (Moses, Psalm 90). It gives one pause for more than a little thought that $\$10^6$ of computing hardware a decade hence may be reasonably expected to bring us to a divine level of information processing capability.

But you might respond that we have no detailed idea of how the brain works, and thus are completely unprepared to provide the software to drive the hardware just scoped out; you might even protest that we do not presently have more than a sketchy, highly incomplete wiring diagram for the human brain's cerebral cortex. My present answer would be that the information content of the entire human genome is only 10^8 bits, and it is difficult to believe a priori that much more than 10^7 bits — 10 percent of the entire set of human genetic information — are devoted to specifying the structure of cortical neurons, their various programs, and their interconnections; after all, instructing the liver's mitochrondria how to tranform fine liquor into a hangover, telling toes how to grow toenails and the cerebellum and the spinal cord how to potentially carry out the real-time control of a ballet dancer all are competing for their share of the 10^8 bits. I would then add that a large physical modelling program of 50,000 lines of FORTRAN expands to a few hundred kilowords of program storage, or about 10^7 bits of reasonably non-redundant information. That a single motivated and intelligent human working for a period of the order of a year can create a structured set of 10^7 bits of globally directed, highly interrelated bits I find to be notable in this context — man-years, not man-centuries, of high quality human effort should suffice to write an engineering specification of the human cerebral cortex's structure and function which would be adequately detailed — and accurate — for emulation programming purposes. Such work may get underway seriously once microneurophysiologists have given us a few more hints, hints that I expect will be forthcoming during a normal course of events in the present decade.

Aren't Those the Opening Bars of the Finale?

I, therefore, am led to the conclusion that two of the highest leverage areas of applied research during the '80s will deal with developing a reasonably high bandwidth, full duplex, time-stable interface between interesting portions of neural nets of the human central nervous system and semiconductor chip 'edges' — bonding pads or whatever — and with creating a reasonably compact, efficient, implantable glucose-oxygen fuel cell, or its functional equivalent, for supplying something like ten watts of average electrical power production within the human body. This conclusion seems compelling simply because the far edge of near-term dealing with the information onslaught will certainly involve at least some humans carrying substantial silicon implants, *mental* protheses with information processing and data storage capabilities which will greatly outstrip our native ones in certain useful — possibly initially highly specialized — respects.

Note that paying even as much as a few femtojoules per logic switching operation and a nanowatt per bit of memory — i.e., using completely unimproved 1981 technology — potentially buys 10^{10} bits of instantaneous recall and multiple CRAY-1s worth of processing capacity in a package comparable to that of a cardiac pacemaker, moreover, one whose glucose and oxygen requirements for power supply would not be much greater than that of the human brain (a dozen electrical watts would suffice).

A decade hence then, human-silicon hybrids may be walking around, looking just like you and me, but each endowed with a thousand times the short-term memory and a ten billion times the specialized data processing power as each of us enjoy today. There is thus sustantial reason to believe that these hybrids may be more "successful" in some of the Darwinian or sociobiological senses than will ordinary humans, with our information processing and storage capabilities unaugmented by silicon implants.

It would be surprising if it did not quickly become feasible for such hybrids to rather frequently re-program their implants so as to conduct an ever larger fraction of the exercising of their organic brains which we call 'consciousness' in the logic sections of these implants, and to store an ever larger fraction of memory-of-events and the results from conscious processing of it which we call 'self' in the memory portions of these implants. The ever greater speed and fidelity of processing and recall which is likely to result as the technology underlying such implants and their programming continues its exponential advance in time will not only strongly stimulate such a transfer of mental functions from a tissue base to a semiconductor one, but clearly will soon alter the nature of such human-silicon hybrids beyond our present ability to imagine.

Of even more profound consequences from our present world-view, perhaps, is the very real transmigration of the human soul from its present seat in the all-too-mortal flesh to a new home in failure-proofed hardware which will thereby occur — not only will the human ego thereby be able to naturally and unambiguously survive the inevitable dissolution of the body which initially sustains it, but it may choose to transcend the flesh just as quickly post-implant as valued memories are recalled from neuronal storage for purposes of being moved into a silicon base. One aspect of near-term human dealing with the information onslaught may thus be physical immortality and perfectability of a hitherto little anticipated nature. The concomitant ability to do high bandwidth transmission of the bit pattern representing one's entire consciousness-and-self around the physical universe at lightspeed on modulated electromagnetic radiation from host hardware at one location to that at another, possibly far-distant one, may prove to be of greater practical, if lesser philosophical, importance.

That human transfiguration of such a profound degree might occur by the close of this century on a technological base only relatively little advanced over the existing one I find to be most striking. I therefore suggest to you in conclusion that only the near-term aspects of dealing with the information onslaught are of human interest, and that the next step is about to be taken in the eons-long procession of species: we will be so altered as a race so soon by coping with the information onslaught that taxonomists a century hence will declare *Homo sapiens sapiens* to be extinct. For the first time in the history of life on this planet, though, the extinction of a species may be voluntary, moreover on an individual-by-individual basis and within a very few generations. Neither artificial nor natural intelligence may be found on Earth a century hence, though intelligence of degrees of which we do not presently dream may then grace the home planet of Man.

Why so quickly? Consider the effective time required to usefully twiddle bits via random mutation and natural selection in a strand of primate DNA as compared to that needed to purposefully alter bits in a computer's memory. A new species of computer-enhanced humans — unable to effectively exchange information with earlier versions or unwilling to make the attempt — can be expected to arise literally overnight, not just once, but as frequently as such entities can devise and agree upon new modes for continuing the exponentiation in time of their information processing capabilities. In brief, 'though we may not all die, we will all be changed, and in a moment, in the twinkling of an eye ....'

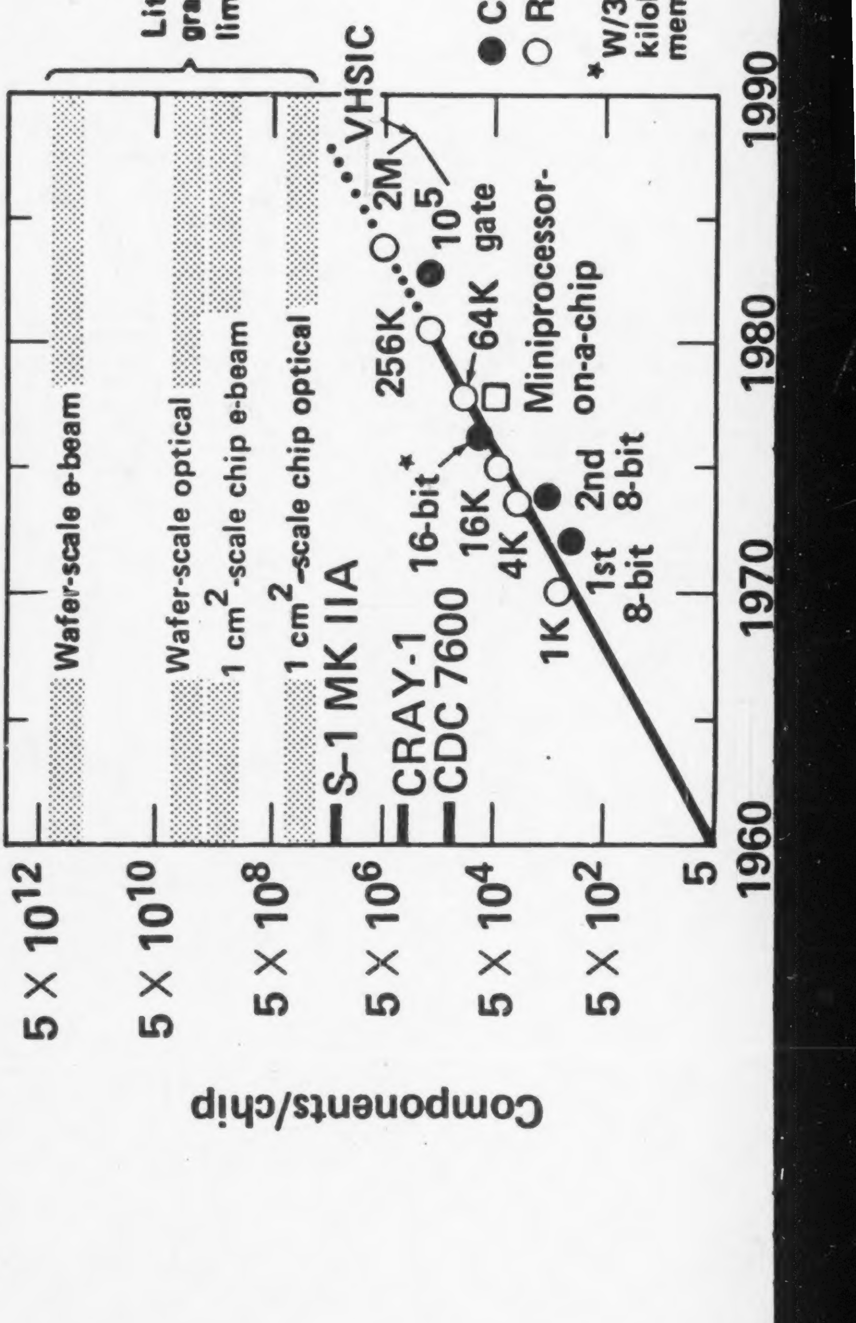
It was more than a third of a century ago, Edward has recalled, during a wartime lunch break at this Laboratory that Enrico Fermi opened the modern discussion of intelligent life elsewhere in the universe by inquiring in a context so empty that his succinct question served very naturally to define the new one, "Where is everybody?" Extensions of the military work in which he and his colleagues were then engaged have so captured the imagination of the intelligentsia since then that it is almost universally assumed that civilizations elsewhere in the universe inevitably attain the capability of nuclear (or perhaps biological) warfare, thereupon to quickly and irrevocably suicide; thus, we never hear from them. Perhaps the 'standard' evolutionary track of intelligent planetary life is somewhat different, and is to be inferred from the not unrelated premises that intelligence strives to process ever greater amounts of information more and more effectively, and that the rise of nuclear weapons technology in a typical civilization occurs at most a short time before the advent of electronic digital computing. I therefore suggest to you that a more valid response to Fermi's inquiry may be, "Pity forbids their appearance locally, for yet a little longer."

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government thereof, and shall not be used for advertising or product endorsement purposes.

FIGURE

DIGITAL LOGIC CAPABILITY COMMERCIALLY AVAILABLE ON A SINGLE CHIP



NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately-owned rights.

Reference to a company or product name does not imply approval or recommendation of the product by the University of California or the U.S. Department of Energy to the exclusion of others that may be suitable.

SYNTHETIC APERTURE RADAR DATA PROCESSING: THE CHALLENGE AND OPPORTUNITY POSED BY ORBITING SYSTEMS

Lowell Wood
Lawrence Livermore National Laboratory
Livermore, CA 94550

The computer-centered synthesis of a radar antenna many kilometers in effective size has been demonstrated in many (unclassified) instances, but in no more spectacular scale than in the NASA/JPL SEASAT-A satellite-borne experiment for ocean surface imaging conducted in 1978. This essay briefly reviews the data processing aspects of the SEASAT-A experiment, and the implications of existing and near-term radar and data processing technologies for enhancing the resolution and area coverage rates of this surveillance technology.

BACKGROUND

There are two notable features of JPL's approach to SAR data processing. The first is that it's strongly oriented to handling the relatively very high area coverage, high bandwidth SAR data from Earth satellites, particularly data streams from SEASAT-A and the Shuttle Imaging Radar (SIR) series. The second is that JPL's SAR data processing hardware is a very innovative, relatively low-budget, ad hoc system, and that all JPL's SAR data processing software is very tightly wrapped around this high-speed memory-starved, multiple array processor-implemented system.

This system, the Interim Data Processing System, processes SEASAT-A data about 600 times more slowly than the data were taken, i.e., its most recent upgrade permits it to synthesize an image of standard format in about 2.5 hours from a data stream transmitted in real time by SEASAT-A over a 15 second interval. Such an image contains about 36,000,000 picture elements (pixels), and is roughly a square with 6,000 pixels on each edge. A pixel at the center of such an image typically represents an area on the Earth's surface of about 16 meters by 18 meters, so that a standard format image obtained by processing a frame of SEASAT-A data represents an approximately square area of the Earth's surface of about 100 km by 100 km size.

The JPL Planetary Radar Program has been kind enough to make available to LLNL not only the published documentation treating their SAR data processing algorithms, but have also provided us with a listing of their operational data processing software and a set of magnetic tapes of representative pre- and post-processed SEASAT-A data; they have also graciously entertained substantial amounts of detailed questions. This has made possible the estimating of the computational labor involved in processing the type of SAR data returned by SEASAT-A with their basic algorithm set, and the mapping of such labor onto LLNL's S-1 Mark IIA supercomputer hardware.

SAR DATA PROCESSING REQUIREMENTS

The processing of SAR data returned during 15 seconds of SEASAT-A flight and corresponding to a nominal 100 km by 100 km imaged area requires about 168,000 Fast Fourier Transform (FFT) operations on data sets of 2048 complex elements each (or their equivalent in FFTs performed on 512, 1024, and 4096 complex element data sets).

Also required are 1,025,000 vector operations (add, multiply, or the equivalent in more complex operations, of which computing 4-dimensional-distance represents nearly 90% of the total arithmetic endeavor) on vectors of 1,024 data items each (or the equivalent in vectors of 512, 2048, or 4096 items). This constitutes the arithmetic effort only; input/output (I/O) and non-arithmetic data manipulation tasks also must be performed.

S-1 SAR DATA PROCESSING CAPABILITIES

Neither the architecture nor the implementation of the now-operational S-1 Mark IIA uniprocessors were optimized for SAR data processing. Nonetheless, the quite general purpose and very high performance microengines in the Mark IIA are reasonably well-adapted for such processing work, as they were designed to very efficiently perform a wide variety of data processing tasks related to DoD's interests. The significant fraction of the hugh high-speed memory capacity of the Mark IIA systems which is actually implemented confers decisive overall performance advantages, relative to the current JPL SAR data processing hardware; the latter seems to be severely limited by the necessity to store virtually all of its data on slowly accessed discs.

We estimate that the execution of the FFT and vector operations corresponding to generating a single SEASAT-A image on the hardware of a single S-1 Mark IIA uniprocessor would require about 110 seconds of wall-clock time. The inputting of 18 megabytes of raw data and the outputting of 36 megabytes of processed data appear to involve 27 seconds of maximum bandwidth data-transferring connection to an IBM 3380-type, 2 megabytes/second disc drive, or about 25% of the execution time of these "kernel" procedures; however, these two operations can proceed concurrently at the end of a processing cycle through the use of two such disc units, so that only about 18 seconds of wall-clock time would be required for (maximally efficient) I/O. Much of this time could probably be further overlapped with processing.

Moving data between the data cache memory within an S-1 Mark IIA uniprocessor and the bulk high-speed memory connected to the processor is calculated to require as much as 20% of the execution time of the kernal operations, although this time expense could be largely eliminated by special microcode to direct the movement of data between the cache and bulk memories (almost) completely concurrently with processing.

About 135-155 seconds (depending on whether ad hoc microcode is utilized) of total wall-clock execution time - - including all overhead activities - - on an S-1 Mark IIA uniprocessor should be required to process 15 seconds of SEASAT-A real-time data into the standard JPL SAR image format. In order to sustain such a processing rate, the uniprocessor would have to have at least 24 million words of high-speed memory attached to it and at least two IBM 3380-level disc drives.

Because each Mark IIA uniprocessor in an S-1 Multiprocessor configuration of any size could process SEASAT-A (or SIR-X) data on a per-frame basis, without any interaction with its fellow uniprocessors, the effective speed relative to real-time at which such a Multiprocessor could handle such data would be about [150/(15*N)], where N is the number of processors in the Multiprocessor configuration and 150 seconds is the time assumed to be required for a single processor to digest a 15 second-duration image frame of SAR data from the satellite. Thus, an S-1 Multiprocessor with 10 member Mark IIA uniprocessors could process such an SAR data stream in real time, while one with 16 such uniprocessors would stand idle close to 40% of the time (or, more likely, would be used to process the SAR data to somewhat higher spatial resolution). However, no data frame would be available in image form for 150 seconds after processing commenced on the corresponding SAR data, even though a new data frame would be available from the Multiprocessor every 150/N seconds, on the average.

The S-1 multiprocessor could also be used so that all N member uniprocessors would be computing in close collaboration on one and the same SAR data set at any given time, so that the latency between start and completion of processing on a given data frame would be reduced roughly N-fold, i.e., would be of the order of 150/N seconds. However, we haven't yet analyzed this more complicated situation in sufficient detail to estimate the actually obtainable latency at all precisely. (This type of processing would presumably be of interest only in an application in which there was a high premium of having SAR data reduced as soon as ever possible after receipt, e.g., in support of battle management.)

The S-1 Mark III and Mark V uniprocessors are estimated to be 2 and 10 times faster respectively than the present Mark IIA prototype systems. Multiprocessors implemented with them would not only be much more compact, but could be employed to process SAR data to the limits of the spatial resolution intrinsically present in it. The SIR-X hardware which JPL hopes to fly in the '87 time frame might have the capability to resolve 10-30 cm scale objects with its K and X band channels, if operated in "spotlighting" mode, and if Mark V-level data processing were available.

SUMMARY

In summary, a single S-1 Mark IIA uniprocessor is estimated to be able to process SEASAT-A SAR data about 60 times faster than JPL's current special-purpose SAT data processing hardware; 15 seconds of SEASAT-A real-time SAR data can apparently be processed in 135-155 seconds of execution of a single s-1 Mark IIA uniprocessor. An S-1 Mark IIA Multiprocessor system with 9-10 processors would apparently be able to process the SEASAT-A SAR data stream into JPL's standard 100 km by 100 km image frames with 16 meter by 18 meter central pixel resolution as fast as the satellite transmitted it. Improvements in processing efficiency may ultimately permit attainment of processing rates 40% higher than these estimates, while presently unforeseen impediments might reduce the average rate to as little as 50% of these estimates. S-1 Mark III and Mark V systems, used in conjunction with currently foreseeable SAR hardware advances, could generate photographic-quality radar images from sateelite-borne systems. All-planet coverage within an hour of when desired at photographic-quality resolution could be provided by NAVSTAR/GPS-scale constellation of SAR satellites interfaced through a somewhat augmented version of NASA's Tracking and Data Relay Satellite System to a suitable S-1 Mark V Multiprocessor system.

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government thereof, and shall not be used for advertising or product endorsement purposes.

Printed in the United States of America
Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161
Price: Printed Copy $; Microfiche $4.50

Page Range	Domestic Price	Page Range	Domestic Price
001-025	$ 7.00	326-350	$ 26.50
026-050	8.50	351-375	28.00
051-075	10.00	376-400	29.50
076-100	11.50	401-426	31.00
101-125	13.00	427-450	32.50
126-150	14.50	451-475	34.00
151-175	16.00	476-500	35.50
176-200	17.50	501-525	37.00
201-225	19.00	526-550	38.50
226-250	20.50	551-575	40.00
251-275	22.00	576-600	41.50
276-300	23.50	601-up[1]	
301-325	25.00		

[1] Add 1.50 for each additional 25 page increment, or portion thereof from 601 pages up.

CONCEPTUAL BASIS FOR DEFENSE APPLICATIONS OF SUPERCOMPUTERS*

L. L. Wood
University of California, Lawrence Livermore National Laboratory
Livermore, California 94550

"What if they gave a war, and no American had to come?" Is such a situation possible and, if so, when and at what cost could it be realized? What is the likelihood that America's adversaries could attain such a capability first? These questions are explored briefly below.

Basic, long-term trends in American society suggest that the useful military power and thus the ability to influence world events of the United States will depend in the future to an ever-smaller degree on Americans under arms. Not only does the political cost of large armed forces continue to climb, but the political toll involved in deploying them in harm's way has become almost unbearably high.

The economic consequences of large American armed forces are nearly as daunting. At the present time, over half of the defense budget of the United States is committed to personnel costs of all kinds. It is not an exaggeration to say that the world's richest nation considers itself hard-pressed to be able to afford to defend itself. The costs of men under arms are thus the single largest contributor to a perceived inability to afford an adequate national defense.

These considerations suggest that the U.S. move to alter its defense posture toward one involving substatially fewer men under arms and far smaller casualty rates and total casualties in the event of hostilities, while simultaneously attaining significantly greater overall force effectiveness.

Is this possible? It is suggested here that it is possible by the aggressive use of battlefield robotics, utilizing integrated state-of-the-art supercomputing digital processing to perform the two key requirements: target acquisition and weapon delivery. With respect to target acquisition, the questions becomes one of the extent to which digital computing can substitute for human effort in locating targets in military environments. Is it possibly feasible to reproduce its function with digital computers in the foreseeable future? Probably not.

However, it's almost certainly not necessary to do so for virtually any military purpose. The human vision system performs a large number of subtasks --e.g., stereopsis, texture classification, inference of shape from shading, color perception (distinct from light frequency analysis), and optical flow field analysis -- which constitute very large processing burden, but which are virtually irrelevant to the highly restricted type of processing needed tc perform target identification and localization. This latter type of work is a very small subset of the computing work of the human visual system, is quite isolated from general intelligence, and is instead embedded in a cognitively quite simple world.

The human visual system, in contrast, has rich interconnections to the rest of the brain, and can be influenced at all but the lowest levels of processing by expectations generated elsewhere within the brain. Small clues in the visual field index into vast stores of knowledge in the brain, which in turn direct visual computations. Contemporary target classifiers/localizers can afford to do without these interactions, as they need only to identify a small class of objects (the targets), and can clump the rest of the scene into the category of "noise" -- they don't need the ability to go from seeing a tusk and a trunk sticking out from behind a tree to hypothesizing an elephant hidden behind it. If they see a tank's cannon barrel poking out of a clump of trees, they do need to make the correct inference, but they don't have to handle giraffes. Since the number of possibilities of interest to them are so small, they just don't need the ability to index into a vast data base that the human visual system requires.

Expert estimates of the digital computational effort required to inspect a high resolution image of a scene and to recognize and precisely position objects of interest within it suggest that the digital computing capability to recognize, classify, and localize the targets of military interest in a real-world scene as rapidly as the scenes are presented by imaging hardware is within the current state-of-the-art of the underlying digital semiconductor technology. These estimates are detailed later in this Appendix.

Even if it is stipulated that it is technically possible to classify and localize targets of military interest purely by digital computing, it isn't clear that it's militarily feasible to do so. Can the required digital computing systems possibly fit within the weight, volume, and power consumption limitations of the military field environment--within the capabilities of a remotely piloted vehicle (RPV), for example--or are they behemoths which can be situated only on a large and sturdy laboratory floor, powered and cooled only by tons of supporting equipment? In Appendix G we present the case that supercomputing in packaging eminently acceptable to the military applications environment is nearly in hand. Upon reflecting that the total mass of active devices and their necessary interconnection of a modern supercomputer is less than an ounce and that the electrical power which must be expended in processing data at the rates of current supercomputers is less than a milliwatt, this assertion becomes credible.

But suppose that the delivery of means for neutralizing enemy targets isn't utterly trivial compared with finding and classifying such targets. Is it still necessary for American-manned aircraft and artillery to dispatch ordnance to such targets, or for American-driven tanks and armored vehicles to physically seize or destroy them, or can high-performance digital computing make a qualitative difference here, as well as with the targeting problem? We discuss below the computing performance required for various levels of robot performance in military environments. It is remarkable that the current estimates of robotics experts of the required performance levels are no greater than those estimated to be required for real-time target localization and classification--of the order of 1000 MIPS of scalar processing. Thus, when 1000 MIPS computing hardware is realized, it seems likely to be highly useful not only in finding military targets, but in neutralizing them, as well.

What are the essential features of a program to attain these levels of computing capabilities, which offer the promise of automating conventional warfare to the point where robot-controlled platforms may even cost-effectively attack strategic weapons systems (e.g., terminal phase intercept of strategic reentry vehicles with "brilliant" high performance rockets carrying non-nuclear kill ordnance)? They appear to be two in number.

The first is a sustained, well-funded, goal-oriented program aimed at the early attainment of wafer-scale integration of very high performance computing systems, for reasons of system radiation hardness and acceleration tolerance, as well as minimum package size and mass. This thrust should feature minimum lithographic linewidths, in order to improve system speed while holding its bulk and power consumption to a minimum, and must emphasize comprehensive redundancy at all levels of implementation, for reasons of both initial yield and in-service durability.

The second is a much more vigorous and focussed program aimed at developing robust algorithms for target recognition, localization, and classification in realistic military environments, over the full class of targets of interest and in the presence of plausible countermeasures, and for controlling the various types of robots of immediate military interest in realistic military environments. This effort should emphasize early and continuing exploration of the effectiveness of multiprocessing for each of the target spotting and neutralization applications. The utility of various approaches can thereby be explored with real hardware at the earliest possible stage of computing hardware development, by replicating and interconnecting slower hardware to stand in for subsequently available higher performance single units.

The infrastructure on which work of this type must be built is at least as advanced in the U.S. as anywhere else. However, it is currently substantially less exploited than is the analogous capabilities in Japan. The fundamental differences between the present American and Japanese postures in this area are ones of funding and of focus. Matching the total Japanese funding levels in these two areas would require funding of the order of $100 M/year in each of them. Attaining comparable focus would require a national policy statement and a significant institutional committment in the Federal Executive Branch (presumably within DOD).

When could significant milestones be set in place in this endeavor? Within a few years of the time a compact, power-efficient supercomputer is available in prototype, a general purpose military robot should attain a similar stage of development, assuming concurrent development of robotics. Special purpose robots, such as silicon pilots for the current generation of cruise missiles and RPV's, may become available somewhat more rapidly. One thousand MIPS monolithic supercomputers could become available in the 1985-87 time frame, if their development becomes a significant national objective in the near future.

Estimates of the Required Performance of Real-Time Target Recognition and Localization

Consider a requirement to recognize and localize all targets of a given type in a high resolution image of a portion of the real-world. For the sake of specificity, take the image to be 1024 x 1024 picture elements (pixels) in each dimension--about twice as good resolution as a standard American TV image--with 8-12 bits per pixel of grayness information. Take the image to be outdoors scenes with arbitrary surface topography and vegetation in which ground vehicles of interest are embedded and whose general orientation relative to the camera is known. Assume that the required level of sophistication in target recognition is roughly that of a seven year old who has been given the opportunity to examine the target objects in some detail, e.g., who will be able to recognize and classify tanks and trucks but may be puzzled by a Trojan Horse.

The digital processing requirements (in millions of operations) required for ad initio recognition and localization of targets under such conditions with contemporary hardware vision technology is as follows:

	Scalar	Vector	
	Ops	Adds	Mults
64M adds, 32M mults for 32x32 convolution, and similarly for 16 x 16, 8x8, and 4x4 convolutions (to provide a reasonable spatial frequency range)		128	64
Similarly for extensive thresholding, thinning and other global array-type computations		128	64
Shape extraction	500		
Recognition/localization	500	___	___
Totals Millions of Operations	1000	256	128

Scalar operations are those not amenable to vector array processing, e.g., regular DEC PDP-10 class instructions, including extensive condition branching. (These estimates are based on what the ACRONYM software uses, which is written in LISP for PDP-10 execution.) The vector adds and mults columns are requirements for addition and multiplication work on arrays, for which vector processing hardware is well-suited. The address space requirements for this processing are 128 megabytes, half of which must be physically implemented with high-speed memory.

These estimates assume complete, ab initio recognition, i.e., processing from scratch, for each image. If the processing from image to image is linked together, then the assumption of object coherence could reduce the 1000 million scalar operations (but not the vector operations) by a factor of 4 to 8, after the first few images were processed. (Implicit in object coherence exploitation is significant commonality between successive images of the terrain being viewed, and a crude estimate of the camera's velocity through this terrain.)

In summary, the framing rate at which humans see the world (less than ten overlapping images perceived per second, limited by the reversal rate of retinal pigments) can be matched by digital hardware which is capable of executing roughly a billion scalar operations and a half-billion array operations per second. The digital hardware will of course not be comparable to the human vision system for any function but that for which it was designed and programmed: the recognition and localization within an imaged scene of target objects of interest.

How might such hardware be implemented within current technological constraints? There appear to be no problems involved in applying on the order of 64 processors to 64 overlapping regions of the image to perform shape extraction. There might be an overall increase as large as a factor of 2 in the number of operations required, due to the necessity of overlap to get locally global shape features. The next stage would be local groupings of shapes: again, as many as 64 processors could be effectively used in this application, and an effective speed-up of at least 32 on the wall clock, relative to a single processor, could also be expected. The final stage in processing is detailed screening and classification of candidate groupings. In this case, at least one processor per candidate group could be employed, so the speed-up attainable with use of multiple processors depends somewhat on how many targets one can expect to see at once, and how many false alarms will get through the local grouping stage. The latter is somewhat terrain-dependent, with pertinent experience suggesting that the average number of false alarms is comparable to the number of targets.

A multiple processor interconnection using of the order of 64 processors, each capable of 16 million scalar operations per second and 6 million vector adds and multiplies, is therefore expected to be capable of localizing and classifying a restricted set of targets in real-time from real-world images. A few hundred contemporary processing chips, supplemented by two thousand state-of-the-art quarter megabit DRAM chips, are apparently sufficient to constitute the hardware needed to implement this vision system with off-the-shelf components.

If smaller numbers of most robust processors were to be used in this application, roughly twenty S-1 Mark IIA uniprocessors, each executing 125 million of the required vector additions and multiplications per second and about 50 million-equivalent of elementary scalar operations, would be sufficient. In this case a reimplementation focussed on a higher scalar to vector execution rate ratio would be indicated.

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government thereof, and shall not be used for advertising or product endorsement purposes.

Printed in the United States of America
Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161
Price: Printed Copy $; Microfiche $4.50

Page Range	Domestic Price	Page Range	Domestic Price
001-025	$ 7.00	326-350	$ 26.50
026-050	8.50	351-375	28.00
051-075	10.00	376-400	29.50
076-100	11.50	401-426	31.00
101-125	13.00	427-450	32.50
126-150	14.50	451-475	34.00
151-175	16.00	476-500	35.50
176-200	17.50	501-525	37.00
201-225	19.00	526-550	38.50
226-250	20.50	551-575	40.00
251-275	22.00	576-600	41.50
276-300	23.50	601-up[1]	
301-325	25.00		

[1] Add 1.50 for each additional 25 page increment, or portion thereof from 601 pages up.

Mrs. **Lloyd**. We will move along since we will need to vacate this room by at least 4 o'clock.

Our next panel includes Dr. Sidney Fernbach from Control Data Corp.; Dr. George Olson, from Colorado State University Research Foundation, and Dr. Robert McCrory, University of Rochester, Laboratory for Laser Energetics.

Thank you, gentlemen, for being with us today.

STATEMENTS OF SIDNEY FERNBACH, CONSULTANT; GEORGE OLSON, COLORADO STATE UNIVERSITY RESEARCH FOUNDATION, FORT COLLINS, COLO.; AND ROBERT McCRORY, UNIVERSITY OF ROCHESTER, LABORATORY FOR LASER ENERGETICS, ROCHESTER, N.Y.

Mrs. **Lloyd**. We have copies of your prepared statements and they will be included complete in the record so you may summarize or proceed as you wish.

Dr. **Fernbach**. Thank you. If I may, I would like to make a correction. I am not representing Control Data Corp. I am a private consultant, and I do consult with Control Data Corp. but I am speaking on behalf of myself.

Mrs. **Lloyd**. Dr. Fernbach, we will correct that for the record.

Dr. **Fernbach**. Thank you.

Mrs. **Lloyd**. Please proceed.

Dr. **Fernbach**. Thank you, Madam Chairman, and members of the committee.

I am very happy to have you invite me to speak about my favorite topic.

For the record, I would also like to make another statement. There has been a great deal of confusion between the supercomputer development that the Japanese have started and the fifth generation computer development.

These are two distinct projects which in this country we tend to lump together. I want to make it clear that MITI is subsidizing or spending about $450 million over the next 10 years for the fifth generation project which is mostly artificial intelligence oriented. The idea is to build computers that can help man think and augment his intellectual capabilities.

These machines that they may build may never have arithmetic units in them, they may never be able to do any arithmetic at all. In other words, they are not useful as numeric supercomputers, the subject about which we are talking today.

The supercomputer project that MITI has sponsored has offered about seven different companies funds to do research in superconducting and other techniques, to develop components as well as new architectures for numeric supercomputers.

In addition to that, of course, the three major companies, Hitachi, Nippon Electric, and Fujitsu, are building their own supercomputers at their own expense.

These are the computers that will be available within the next few years. A MITI effort will not result in supercomputers until the 1990's.

I would like to express some of my thoughts on supercomputers. You have my written statement and so I will deviate from that.

Mrs. Lloyd. Please proceed as you wish, Dr. Fernbach.

Dr. Fernbach. Thank you.

One point that has never been brought out very clearly is that uniprocessors, that is, processors that are sequential, doing one step at a time, are reaching the limits of their capability. This is the type of processor that IBM builds in a very large number and the number installed is counted in the thousands.

However, if you look at some of the installations you will find tens of those machines working together. The reason that they have tens of those machines is that there is no supercomputer built by IBM that would help them do their job.

They dare not go to the other manufacturers because they are scared of the scientific supercomputers that exist today. These are difficult to use. The software is often lacking. And even within the scientific communities, these are not used very well. As a matter of fact, one could improve performance on those machines by an order of magnitude if more research were done.

And so with the uniprocessors reaching their limit, what is the world going to do about increasing their computer capabilities in the future? They are going to have to use supercomputers in the future.

I would argue that the market for supercomputers will become very large over the next 10 years. But this will not happen unless we support the development of the software, the algorithms, the numerical techniques that will make them easily useful to potential users.

This, I think the Japanese are going to try to do. And we should put a lot more money into research to make this happen here.

This might be a good task, incidentally, that is, the research, for the laboratories and the universities. They have tremendous intellectual capability at these organizations and that is the only source for such capabilities.

So I would argue that there are ought to be an extensive amount of research done in these areas.

Another point I would like to make very strongly is that the market could be very broad. In other words, the applications could be broad. These supercomputers can be used in banking, insurance, in all of the areas where commercially available IBM machines are now being used, in addition to the scientific areas.

The first problem that one faces is how do you get access to these computers to do your research? And this is where the Lax report has pointed out severe deficiencies in this country. It seems to me that we must make access available to the universities where the capability resides to provide the algorithms and so on.

The question was raised earlier as to how many might be useful. I would say that we have to start immediately with about a half dozen supercomputers.

I would like to suggest that they not be placed at universities. Universities, in all of the 30 years that computers have been in use, have never demonstrated a capability for operating a computer center adequately.

So I feel that the universities ought to get together and form consortia to provide an NCAR type of environment; perhaps the National Science Foundation or Department of Energy could set up an institute of computer science research, perhaps a number of them, operated in very much the same way that NCAR is operated, so that access can be made available to the researchers at the universities and even some of the other laboratories.

I also would suggest that one not use existing supercomputers to any great extent. I don't believe, first of all, that they are available. But, second, I do believe that it would be a much better decision to spread the capability throughout the country.

I would like to see a number of laboratories set up regionally so that we can train new people in the art of maintaining computers, supporting computers, and providing the services necessary to make these work well.

Again, the medium might be through setting up institutes by way of NASA, DOE, or NSF.

By spreading them out we may also succeed in lowering communications costs. These computers should all be interlinked through a network so that users can have access not only to a regional one but to another remotely located.

The reason for this is that it is very difficult for one site to have expertise, or develop expertise, in all disciplines and make those disciplines available on a single set of computers. It would be far better, perhaps, to have an MFE network located at Livermore, a chemistry network located at Carnegie-Mellon or some such place, so that one can develop expert capabilities in specific areas at different sites.

I would like to point out that there is a good deal of research, perhaps more these days than ever before going on in supercomputer components and in supercomputer development at the universities.

Many universities are exploring the possibilities of putting together small chips, or smaller computers into arrays. The direction we seem to be taking is that of going into vector processing or multiprocessing. Multiprocessing consists of having a large number of computers operating together, simultaneously, on a single job. We don't know how to do that well, at present.

There is very little, if any, research going on in the algorithms needed to design the mathematical problems, the codes, the operating systems, the compilers, to allow us to efficiently use a multiprocessor system.

However, there is a lot of research going on in putting hardware together. It is a very uncoordinated activity and I would like to see more coordination through, perhaps, the National Science Foundation, because I do believe we are wasting lots of money. If we had more coordinated activity, we would be able to make much more progress in developing new architectural designs for the future. This, I think, is very important.

There is a great advantage in having universities pursue these activities because they also provide the manpower which would be available for developing these supercomputers for manufacturers in the future. In addition there is a great deal of fallout from their research projects.

On the other hand, there is natural inclination to build complete systems at universities. I don't believe that that is really the direction we want to take.

Back in the 1950's, many universities built systems. True enough, they did develop many people; that is created Ph D.'s in computer science and computer technology who went on to become the Crays and others who are building the supercomputers of today.

So from that point of view, it is very good. They also may come up with some good ideas that can be incorporated by the supercomputer manufacturers.

However, there are many disadvantages: First of all, they build systems that are one of a kind. To support such a system costs a tremendous amount of money. The software itself would be much more expensive than the construction of the computer. If we put the computer burden on the manufacturer, he could provide software which could be shared by everybody else.

Scientists are also not very inclined to use a machine which is a one-of-a-kind. They like to exchange programs, they like to be able to talk to each other, and they like to use something that is already there rather than to develop from scratch some of the algorithms that they have to use to do their problems.

These universities are also doing some very interesting work in areas such as data flow concepts, in developing new switches, and in using array processors tied to small computers to gain performance. Many of these activities, I think, will lead to eventual applications in the supercomputers that manufacturers will produce.

One of the questions raised was how can Government and industry interact profitably to help us in arriving at our goal to be No. 1 in the supercomputer business?

Industry, first, doesn't like to be told what to do; it likes to go ahead and build its own thing. Mr. Cray won't build whatever you want him to build; he will build what he wants to build.

Industry wants to protect patents so they don't want to get in bed with the designers at the laboratories or universities.

But industry does need support. Despite what some of the industry people told you yesterday, they do need the support. They are reluctant to have the support, I think, because it means new competitors might arise and each one feels that there is a very limited market for himself and doesn't want to share with too many others.

I think it is important for this country to have a larger market and a larger number of industries involved.

It takes about $100 to $150 million to design and develop a supercomputer. That has to be shared somehow.

My notion of how that should be done is by having DOE, NASA, NSF, or whoever may have needs within the Government, order advanced state-of-the-art computer systems every 5 years. In other words, set goals for ourselves, like a 10 billion floating point operations per second system machine by 1990, and order three from each of the manufacturers. A manufacturer doesn't like to be tied too closely to specs, just set the goal, and let him go at it.

But insure him, as the Japanese do, that there will be a market. In other words, order a dozen of them. We will have lots of places to put them: universities, university consortia, or at the laboratories, NASA, DOE. And despite what you heard yesterday, the Department of Defense does have a tremendous use for the large-scale computers. I can give you examples if you care to have them.

One of the areas where industry does fall down is in the development of software. Apparently nobody in the world knows how to develop software well. However, we have a possibility now of using expert systems, or automated systems to develop software. We use automated systems to develop hardware. We should be able to develop techniques to develop software as well.

I would like to see a good sum of money spent on automating software development because that is going to be the most difficult thing for us in the future if we want to spread the use of supercomputers wherever they are needed.

The Government is the only source of funding that might prevail in this area. So I feel that DOE and NSF should be allowed to pursue this with vigor.

Let me just mention in closing my view of the future. I feel that 10 or 20 years from now this world will be dominated by the use of supercomputers. We will have thousands of them. And that the small 3081's will be on your desk. In other words, the uniprocessors that are being manufactured by IBM today will be your new terminals. We will be using them to explore and to do some research, but the large-scale research, which is absolutely essential for this country, will be done on the supercomputers.

Thank you.

[The prepared statement of Dr. Fernbach follows:]

TESTIMONY TO BE GIVEN AT A JOINT HEARING BEFORE THE SUBCOMMITTEE ON ENERGY DEVELOPMENT AND APPLICATIONS

June 15, 1983

By

S. FERNBACH
CONSULTANT

I wish to thank you for allowing me to make some comments on the supercomputer situation in the U.S. and what we might do to ensure that the U.S. maintains its leadership in the development and utilization of this important tool.

The Lax Committee report indicated the need for supercomputers in the research environment of industry, national labs, and the universities. It also emphasized the importance of providing access for researchers who currently have no way to gain access to large scale systems. A third point, which to me is most important and was left to further study, concerned methods for accelerating the development of supercomputers into the 21st Century.

My first comment concerns itself with the market for supercomputers. There are already over 60 Class VI machines (the modern supercomputer) installed. During the 70's there were but a dozen Class V machines built and installed. Before that there were about 100 Class IV supercomputers installed (in the early 70's). I am using the Lax Committee definition of supercomputers; namely that they are the most capable systems in existence at any given time. Compared to the current vector/multiprocessor system, the earlier Class IV computers were relatively easy to program and use. The Class V paved the way for the newer architecture. Even today, however, 7 years after the first Cray-1 was put into operation there has been very little research into optimal ways to use the new machines. The result is that these machines are poorly utilized. We need much more research in these areas than we are presently funding. If access to such machines were made readily available to university and national laboratory researchers, great improvements could be realized. It is primarily because of the relative difficulty of use that the market for Class VI computers is not soaring. We are approaching the end of an era, in which uniprocessors (sequential machines) are close to their limits of performance using current silicon technology. We must start using the more complex systems to match the ever increasing need for performance, even in the general market place. Many commercial and scientific facilities are saturated, some using multiple uniprocessor systems. They must eventually make use of the more exotic vector/multiprocessors, but we need the results of software research and resulting new tools to help these average users overcome their fear of supercomputers and help make these systems easy to apply to any task.

For this we must first furnish access to the researchers and teachers who can develop the appropriate techniques for using these machines.

They will have to develop new mathematical concepts for vector or parallel processing, new algorithms for efficient use, new software to relieve the user of the burdens of understanding the new architecture. This can be done by providing funds to acquire and install more supercomputer capability throughout the country and by providing a network to interlink them in order to make them readily available to the research and software development community. All supercomputer systems of United States manufacture should be in this network.

These new installations could be most easily handled by increasing the capabilities at the national labs. This could allow already trained personnel to provide their know-how to make early access feasible. Machines could also be placed in a number of university environments. This bothers me, first, because I believe that universities should not be in the computer brokerage business, and secondly, because very few universities have demonstrated great ability in running computer centers during the past 30 years. One way of getting around this hurdle is to have a number of universities get together and contract with an outside party to operate a facility for them. This third party could be a national laboratory.

One example of such an operation is the National Center for Atmosphere Research (NCAR). It is funded by NSF and operates through UCAR, consisting of a group of some 50 universities. Some of the national laboratories such as Brookhaven and Fermilab are operated in a similar way. As a matter of fact, it might be well to have NSF through its Math and Computer Science Directorate support a number of Computer Science Institutes at which high level research in computing, applications, and computer science could be supported. DOE, itself through the office of Energy Research could also support a number of sites, similar to its support of the Magnetic Fusion Energy Computer Center.

The National labs already have, for the most part, well developed computer centers and skilled personnel for manning them. If these labs are enlarged to take care of the required research, the overall costs might be lowered but little or no expertise is transferred out to other facilities in the country. I would favor the development of new facilities in various regions of the country in an attempt to develop more talent and capability throughout the U.S. in this important activity. This might also have the effect of reducing communication costs which seem to get higher all the time. I believe that a half dozen new sites might be very welcome.

The importance of new, even more advanced supercomputers must not be overlooked. Left to their own devices, U.S. manufacturers of supercomputers will not move very quickly into new, improved designs. Their perception of the market is based on past appearances. This implies a very restricted number of systems and limited profits, if any. Futhermore, there is no reason to come out with an enhanced system, which might impact the sales of an earlier system. The goals manufacturers set for themselves are usually well within the state-of-the-art. Competitive forces have not yet been so great that they will go all out to be number 1. The Japanese competition may change this. On the other hand, the Japanese competition might just force them out of the marketplace completely. It is imperative from a military, economic, and scientific standpoint for the U.S. not only to stay in the supercomputer business but to be at the very peak at all times.

How can the government assist in keeping the U.S. strong in supercomputer development? DARPA, the largest government funding agency in the area of computer science and computing, concentrates on military related research. There is fallout into the general computing area from these efforts, but we cannot rely on this being the case at all times. Furthermore, DARPA is seeking to develop systems making use of artificial intelligence (AI) techniques that will become important to the U.S. and the military in particular. It is highly unlikely that the DARPA program will contribute significantly to advances in scientific supercomputers. For the civilian market, we must develop techniques and systems to do our numeric data processing at the ever increasing pace needed. This can be done more effectively through NSF and the major using agencies such as DOE, NASA, and NOAA. Funding required for continued development of the scientific supercomputers should be at a minimum of $40 Million per year.

It has been and is being suggested that national labs and universities make their own contributions to hardware development for supercomputers. This is now being done in an unprecedented way. There is a great deal of activity in researching new architectures, primarily at universities, but there is little or no coordination. This has the effect of spreading our manpower too thinly and at the same time increasing overall costs of such research. Should a national lab or university build computers? Unless it be a very special purpose device to solve a particular problem, I say no. This sort of activity was encouraged in the early 1950's. Little or no direct contributions were made to the computer industry through these efforts. The major result (an important one, at that) was the education of a substantial number of computer architects, many of whom later made their mark in the industry or academia. Such can again happen through the new revolutionary potential of microelectronics. Hence, research in these areas should be supported, but better coordination of activities might provide better results. We should not have to depend on products, such as complete systems being developed through these research efforts. Industry should provide the systems, making use of whatever useful research comes out of the universities. It should also be pointed out that the cost of software developed for a single system at a university or laboratory would be prohibitive. If the software costs can be spread over a large number of systems, as industry usually does, then the individual costs become more reasonable.

To provide the most effective support for the furtherance of supercomputer development in this country, the U.S. Government should fund research in component technology, such as GaAs or other materials that might provide higher density, lower power chips than now available. It should also fund at a higher level of support research in mathematics, numeric algorithms, and software needed for the newer architecture now being developed. It should provide industry with new goals every five years, these being such that would promote at least an order of magnitude improvement in performance in this time period. Between NASA and DOE, the utilization of supercomputers today is of the order of 20% to 30% of the industry's production. These agencies should be the ones selected to provide model sites for the newly developed systems. Perhaps as many as 10 computer systems just beyond the state-of-the-art can be ordered every 5 years to create the proper environment for a prosperous industry. Incidentally these should be complete systems, including all the I/O required, mass storage units, network, terminals and most importantly, the software.

Because a great deal of know-how in software design and network system architecture resides in the national laboratories, it might be well to have industry and government join forces, the manufacturer supplying the hardware, and the labs (singly or jointly) providing both the software and systems integration. If these efforts are carried out in parallel, each generation of machines should be ready to use at the end of the 5-year construction period. The experiences gained at the ten sites could be fed back to the manufacturers so that they may have some impact on the succeeding machines.

Incidentally, the ten machines should not all come from a single manufacturer. Each manufacturer should be given the same goals with the assurance that the government would buy a number of systems for DOE and NASA from each. A best efforts approach should be used. The cost of such a program could be on the order of $150 million over each 5 year period.

E0227J

Mrs. Lloyd. Thank you very much, Dr. Fernbach.

Dr. Olson. We are looking forward to hearing from you now.

[The biographical sketch of Dr. Olson follows:]

BIOGRAPHICAL DATA

GEORGE G. OLSON
1909 Mohawk
Fort Collins, Colorado 80525

Present Position	:	Vice President for Research Colorado State University
Experience	:	16 years university research administration 15 years research director in private business
Education	:	Ph.D. (Physical Chemistry), University of Colorado
Military Service	:	Diesel Engineering Officer, U.S. Navy, 1943-46
Professional Organizations	:	American Chemical Society Society of Sigma XI American Association for the Advancement of Science, Fellow American Institute of Mining, Metallurgical and Petroleum Engineers Licensing Executives Society

Dr. Olson. Madam Chairman and members of the committee: I appreciate the opportunity of coming here and testifying——

Mrs. Lloyd. Excuse me. If you will get a little closer to the microphone, or turn it on. Is it turned on? Thank you.

Dr. Olson. Colorado State University, together with its affiliated research foundation, is the only agency in the world to own or operate both supercomputers, a Cray and a Cyber 205.

We are here to share with you some of our insights into their problems of operation in providing access to the university community.

One of the basic requirements today for the state-of-the-art research in virtually all fields is access to supercomputers. American decline of preeminence in many fields of academic research can be directly related to the lack of availability of state-of-the-art computing on university campuses.

As an example, no American university ever acquired a CDC 7600 even though it remained the top scientific computer from the time it was introduced in 1966 until the advent of the supercomputer in 1977.

Yet, during that time a number of universities in Europe acquired 7600 computers and a number of them have already replaced them with the state-of-the-art supercomputer.

The fact that these computers are not available to American university researchers, not only limits the quality of their research but means they are not training graduate students in writing modern codes, they are not exploring new computational methods, and are actually falling behind their counterparts in Western Europe and Japan.

The capital costs of a supercomputer facility is in the range of $7 to $10 million depending on the size, ancillary equipment, and so on.

It is costs of this order of magnitude that have prevented universities from acquiring state-of-the-art scientific equipment for the past 16 or 17 years.

Colorado State University, through its research foundation, has addressed this national problem through its program of self-financing of scientific equipment.

In 1977, the foundation used the proceeds from an industrial development bond to acquire the Cray-I computer, which it leases to the National Center for Atmospheric Research in Boulder.

In 1982, the foundation installed a 2-million-word, 2-vector pipeline Cyber 205 computer at Colorado State University. The Cyber 205 was installed to serve the following four groups of users:

First, the research program at Colorado State University.

Second, a joint university/industry institute for computational studies in petroleum.

Third, Government laboratories and agencies which did not have supercomputers, such as the National Bureau of Standards, with whom we have a direct telephone tie.

And a national university supercomputer network.

Our studies indicated that it was not likely to find that any one university will have sufficient funding in computational research projects to justify the acquisition and operation of a supercomputer.

It was necessary, therefore, to consider ways in which access could be made to a number of universities. The most successful model of such a supercomputer network is the satellite-based communications network operated by the magnetic fusion energy group at Lawrence Livermore Lab.

CSU is in the process of developing a similar network around its Cyber 205. Its goal is to have 6 to 10 regional university nodes geographically distributed around the country.

Each university operating as a node would be required to serve scientists and engineers from other colleges and universities in the region with both access to the supercomputer and assistance in developing codes suitable to be run on it.

A university-based supercomputer must, by its very nature, serve a great variety of disciplines. It must, therefore, have a strong service group working with it to provide computational and programing support to a variety of users.

In April, the first node of the CSU supercomputer network was dedicated at the University of Massachusetts and it is presently serving New England colleges and universities through the Dartmouth Kiewit Network.

Initially our communication is through leased telephone lines, but as the demand increases we will convert to either a satellite link or the new 64-kilobit telephone line.

It is our goal that as the utilization of the supercomputer at CSU increases, a second supercomputer will be placed at one of the other university nodes.

It is our belief that access to supercomputers must be provided to university scientists and engineers if our country is to maintain its world leadership in basic research and the development of computational equipment and software.

Not only must university scientists have access to state-of-the-art computers, but some of these computers must be located on university campuses.

We may accomplish high-speed computational results at the national labs or at foreign universities, but it is only on our own campuses that we train graduate students, stimulate innovative approaches to problem-solving and provide the freedom of inquiry that lets a professor try an unconventional and, generally, unsuccessful new idea.

There are now four supercomputers serving the higher education community: The Cray-I at the National Center for Atmospheric Research, a Cray-I at the University of Minnesota, a Cyber 205 at Purdue, and the Cyber 205 at Colorado State University.

The Cray-I at NCAR is full to the point where access must be limited. The others are not utilized enough to cover break-even costs.

An examination of the reason for this will give some guidance to the role of the Department of Energy in supporting supercomputer access at universities.

The Cray-I at NCAR is limited to problems in the atmospheric sciences and provides computing support to atmospheric scientists free of charge through its annual budget from the National Science Foundation.

Furthermore, the Cray-I followed a CDC 7600 in the computer center at NCAR that has provided the only academic discipline, atmospheric science, with access to state-of-the-art computational equipment for the past 17 years.

Also, the Cray was installed in 1977, and atmospheric scientists have had 6 years now in which to develop vector codes and algorithms.

To all of the other academic disciplines on which the Department of Energy's basic research program depends, access to supercomputers was first provided last year.

The scientists in these disciplines were totally unprepared for modern computational methods. Their computer programs were not written to take advantage of vector processing capability of a supercomputer and, more importantly, their research project budgets did not contain funds that would allow them access to computers costing a thousand dollars an hour or more to operate.

The Department of Energy could enhance its own academic research program by providing the scientists which it supports with access to one of the existing university supercomputers. This access could best be provided by acquiring blocks of supercomputer time and allocating that time to its regular contractors and grantees.

I cannot urge the members of this committee too strongly to consider some means of access of supercomputers to the academic scientific programs which the Department of Energy already supports. And at the same time, strengthen the computational capability of university computer centers.

Thank you.

[The prepared statement of Dr. Olson follows:]

Statement of
Dr. George G. Olson
Vice President for Research
Colorado State University

Before the
Subcommittee on Energy Development and Applications
and the
Subcommittee on Energy Research and Development
U.S. House of Representatives
June 15, 1983

Chairpersons and Members of the Committees:

Colorado State University together with its affiliated organization, the CSU Research Foundation, is the only agency in the world today to own and/or operate both of the present supercomputers, a Cray-I and a Cyber 205. It is because of our direct involvement with the present day supercomputers that we are in a position to give some insights to their operation, uses and future development from the university point of view.

Today's supercomputers are generally defined as computational devices which will perform arithmetic functions, add or multiply, at rates in the hundreds of millions of floating point operations per second. Previous computers were scaler in their operation and speeds were measured in millions of instructions per second. The supercomputer gets its increased speed not just from the computer architecture but also from the ability to program long "vectors" or sequential operations that may be conducted in parallel processing units or pipelines. The supercomputer then is not just a more powerful computer in memory size and speed of operation but also represents a new type of computing requiring new software, algorithms and codes to be fully utilized.

One of the basic requirements today for state-of-the-art capability in most fields of research is access to advanced computers. The decline of

American preeminence in many fields of academic research can be directly related to the lack of availability of state-of-the-art computers on university campuses. For example, no American university ever acquired a CDC 7600 computer although it was the top scientific computer from the time it was introduced in 1966 until the advent of the supercomputer in 1977. Yet, during that time, a number of universities in Europe acquired CDC 7600 computers and a number of them have already acquired the current state-of-the-art supercomputer, either a Cray-I or a Cyber 205. The fact that these computers are not available to American university researchers not only limits the quality of their research, but means that they are not training graduate students in writing modern codes, they are not exploring new computational methods and are actually falling behind their counterparts in Western Europe and Japan in the basic elements of computer science.

The capital cost of a supercomputer facility is in the range of $7 to $10 million depending on the size of computer memory, ancillary equipment, power sources, etc. It is costs of this order of magnitude that have prevented universities from acquiring state-of-the-art scientific computers for the past 16 or 17 years. It is not likely that such funds will become available from federal sources for scientific equipment acquisitions.

Colorado State University, through its Research Foundation, has addressed this national problem in its program of debt financing of scientific equipment. In 1977 the Foundation used the proceeds from an Industrial Development Bond issue to acquire a Cray-I computer which it leases to the National Center for Atmospheric Research. In 1982 the Foundaton installed a two-million word, two vector pipeline Cyber 205 computer at Colorado State University. The Cyber 205 was installed to serve the following four groups of users:

1. The research program at CSU;
2. A joint University/Industry Institute for Computational Studies in Petroleum;
3. Government laboratories and agencies which do not have access to supercomputers such as the National Bureau of Standards; and
4. A national university supercomputer network.

Our studies have indicated that it is not likely that any one university will have sufficient funding from computational research projects to justify the acquisition and operation of a supercomputer. It is, therefore, necessary to consider ways in which access to a supercomputer on a university campus can be made available to other institutions. The most successful model of a supercomputer network serving both national laboratories and universities is the satellite-based communications network operated by the Magnetic Fusion Energy group at Lawrence Livermore Laboratory.

Colorado State University is in the process of developing a similar network around its Cyber 205. Its goal is to have six to ten regional university nodes geographically distributed around the country. Each university operating as a node would be required to serve scientists and engineers from other colleges and universities in the region with both access to the supercomputer and assistance in developing codes suitable for a supercomputer. A university-based supercomputer must, by its very nature, serve a great variety of disciplines. It must, therefore, have a strong service group working with it to provide computational and programming support to a variety of users. In April, the first node of the CSU supercomputer network was dedicated at the University of Massachusetts and is serving New England colleges and universities through the Dartmouth Kiewit Network. Initially, communication links will be leased telephone lines, but as the demand for access increases, conversion to a satellite link will be made. It is our goal that as the utilization of the supercomputer at CSU increases, a second supercomputer will be installed on the network at one of the other university nodes.

It is our belief that access to supercomputers must be provided to university scientists and engineers if our country is to maintain world leadership in basic research and the development of computational equipment and software. Not only must university scientists have access to state-of-the-art computers, but some of the computers must be located on university campuses. We may accomplish high speed computational results at the national laboratories or at foreign universities, but it is only on our own campuses that we train graduate students, stimulate innovative approaches to problem solving and provide the freedom of inquiry that lets a professor try an unconventional (and generally, unsuccessful) new idea. But that's the nature of basic research, and for too many years now this country has disenfranchised its academic researchers from modern computational resources.

The supercomputer is a special purpose computational machine that, if properly used, could greatly enhance the basic and applied research programs supported by the Department of Energy at universities. The ability to handle the large numbers of calculations of a numerical model of a real physical event and simulate changes on the computer has wide application in energy research. These include the following:

1. Three dimensional seismic exploration;
2. Enhanced management of oil and gas reservoirs;
3. Modeling of secondary and tertiary oil and gas recovery techniques;

4. Modeling in situ combustion of coal;
5. Multivariant studies of the combustion process;
6. River basin modeling and hydroelectric power simulation;
7. Nuclear reactor design and reactor core modeling;
8. Complex power grid modeling;
9. In situ leaching of uranium ores; and
10. Any process lending itself to complete numerical analyses.

In addition, the supercomputer supports such basic science areas as fluid dynamics, aerodynamics, chemistry of large molecules, high energy physics, etc.

The need for supercomputer computational support in Physics has been established by the NSF sponsored report, "Prospectus for Computational Physics", and the National Research Council has published a computational needs analysis for both oceanography and fluid mechanics. In addition, the interagency committee chaired by Dr. Peter Lax stated as one of its specific recommendations the establishment of regional supercomputer centers at selected universities.

There are now four supercomputers serving the higher education community: the Cray-I at the National Center for Atmospheric Research (NCAR), a Cray-I at the University of Minnesota, a Cyber 205 at Purdue University and the Cyber 205 at Colorado State University. The Cray-I at NCAR is full to the point where access must be limited, and the others are not utilized enough to even cover operating costs. An examination of the reason for this will give some guidance to the role of the Department of Energy in supporting supercomputer access at universities. The Cray-I at NCAR is limited to problems in the atmospheric sciences and provides computing support to atmospheric scientists free of charge through its annual budget from the National Science Foundation. Furthermore, the Cray-I followed a CDC 7600 in a computer center that has provided the only academic discipline, atmospheric science, with access to state-of-the-art computational equipment for the past 17 years. Also, the Cray-I was installed in 1977, and atmospheric scientists have had six years now in which to develop vector codes and algorithms. To all of the other academic disciplines on which the Department of Energy's basic research program depends, access to supercomputers was first provided last year. The scientists in all these disciplines were totally unprepared for modern computational methods. Their computer programs were not written to take advantage of the vector processing capability of a supercomputer and, more importantly, their research project budgets did not contain the funds that would allow them to access computers costing a thousand dollars an hour or more to operate.

The Department of Energy then could enhance its own academic research program by providing the scientists which it supports with access to one of the existing university supercomputers. This access could best be provided by acquiring blocks of supercomputer time and allocating that time to its research contractors and grantees. This method of providing access is cost effective in two ways: first, it avoids the overhead costs associated with the grantee institution for a service provided elsewhere; and, second, it assures that funds provided for computation will actually be used for that purpose and not become a contingency budget item.

I cannot urge these committees or the Department of Energy too strongly to consider some means to provide access to supercomputers for the academic scientific programs you already support and, at the same time, to strengthen the computational capability of university computer centers. As the existing university supercomputers become saturated, the higher education community will respond by providing additional machines if it is assured of continuing utilization.

A stronger computing capability within the higher education community will also enhance academic research in the field of computer science. Due to the limited availability and application of the present supercomputers, we have hardly begun to develop the operating systems, compilers, applications software, etc., that will allow us to take full advantage of the existing architecture. At the present rate of progress, we are at least ten years away from having all of the software available to fully utilize the existing supercomputers. Now the Japanese have presented us with a challenge that they intend to build a supercomputer 100 times faster than either the Cray or Cyber 205.

American industry has responded to this challenge with the organization of a cooperative research endeavor under the Microchip and Computer Corporation. American universities would like to assist but to do so must have access to the proper tools. These tools include supercomputers which must be used in design, development and engineering of VLSI chips, components and the next generation computer itself. Nowhere is Computer-Aided-Design or Computer-Aided-Engineering more critical than in the design of the next generation of supercomputers.

The role of the university in the development of the next generation of supercomputers, however, will certainly be secondary to the efforts of the computer industry itself. The complexity and capital requirements of computer development exceed the capability of the university community. The university contribution will properly and importantly be in the basic sciences, the properties of condensed matter, material sciences, physics and chemistry of surfaces, etc. Additional assistance will come in circuit design, algorithm development, data set management, languages, computer graphics and general software development. Without the basic physical and computer science activities of the universities, there will not be the base upon which America's high technology industry can build.

The Department of Energy can play an important role in encouraging the development of future computers. As in the past through the national laboratories it administers, the Department has provided a major market for new computers. Without the potential market of the national laboratories and the Department of Defense, it is doubtful that any company would undertake the expense of developing the next generation computer. It is quite possible that even now a program of joint industry/government funding will be required to match the efforts of Japan in achieving the next generation supercomputer. Perhaps the most important contribution the Department can continue to make is in the development of operating systems and applications software for future computer systems.

In summary, let me state that the Department of Energy is in a unique position to assist the university community in obtaining access to supercomputers. The Department has in place a contract and grant program in basic sciences with the university community. It has the capability of acquiring blocks of time on existing and future university supercomputers and making that time available to scientists it is already supporting. It also has the capability of making excess time at national laboratory supercomputers available to the university community. It should set a goal of assisting in the establishment of supercomputer centers at selected university sites to foster university development of advanced computational devices and software.

Finally, it should use its well established need for even more advanced computational equipment to spur the development of new computers by the private sector.

In closing, let me urge the members of these two subcommittees to carefully consider the endorsement of H.R. 2118 which would greatly simplify the self-financing of supercomputers and other scientific equipment by the universities.

Mrs. LLOYD. Thank you, Dr. Olson.

Dr. McCrory?

[The biographical sketch of Dr. McCrory follows:]

ROBERT L. MCCRORY

Robert L. McCrory is director of the Laboratory for Laser Energetics at the University of Rochester. He is also the founder and chairman of the northeast Consortium of Universities Concerned About Campus Computing. This organization, established in 1982, represents fourteen universities.

He holds a B.S. degree in physics and a Ph.D. in applied plasma physics, both from the Massachusetts Institute of Technology where he was an Alfred P. Sloan National Scholar and an Atomic Energy Commission Special Fellow. Dr. McCrory joined the Laboratory in 1976 as a scientist and leader of the Theory and Computation Group and was named director of the Theoretical Division in 1979. He has also served as associate professor of physics and of mechanical engineering at the University. He is a consultant to several private and governmental organizations, including the Los Alamos National Laboratory, where he was a staff physicist in the Theoretical and Theoretical Design Divisions before coming to Rochester, and to the Lawrence Livermore National Laboratory, the Goddard Laboratory for Atmospheric Sciences, the Air Force Weapons Laboratory, and the Johns Hopkins Applied Physics Laboratories.

Dr. McCrory's major research publications have concentrated on the computational hydrodynamics and hydrodynamic stability of inertial confinement fusion systems. He is the author or coauthor of numerous scientific articles and he has presented papers at scientific conferences in the United States, Canada, France, Germany, the United Kingdom, the USSR, and Japan.

Mailing Address: Laboratory for Laser Energetics
University of Rochester
250 East River Road
Rochester, New York 15623
(716) 275-4973

Dr. McCrory. Madam Chairman, with your permission, I would like to briefly summarize.

In the earliest days of computing, universities provided some of the most advanced computing power and support to their students and faculties. The situation has deteriorated very rapidly over the last two decades.

I think it is a shame that of the 42 supercomputers presently in the United States, one finds only four at universities, in comparison to the larger number of American-made supercomputers which are in foreign universities.

Madam Chairman, and distinguished members, I propose that the United States initiate a program comprised of several parts.

First, I urge that a program of limited duration be initiated that will cost the Nation $90 to $100 million over the next 5 years to fund five new supercomputing sites.

I suggest that the U.S. Department of Energy is, in fact, the correct Agency to take this initiative.

Second, I propose that the Department of Energy, along with other agencies, immediately enhance the capabilities at existing supercomputing sites in order to provide an expedient solution to the problem of access to supercomputers for universities.

My own research has been in energy research in inertial fusion. At present, I am the director of the Laboratory for Laser Energetics, which has invested approximately $100 million since 1972 in inertial fusion research, less than half of which has come from the Federal Government.

I am also the chairman of the Consortium of Universities Concerned about Campus Computing. This is an informal consortium founded last June. The history of that organization is contained in my formal testimony.

The common thread that binds that group is one of access, immediate access. In talking with the various consortium members, the one thing that is clear is that some combination of both Government and university support is going to be necessary to accelerate progress and address the problem of access. This combination is what I shall call the condominium approach to computer financing.

There is an institutional apprehension that some of the things that happened in the sixties and seventies could happen again. And what were those?

One, declining Government support of main frame facilities. Two, at about the same time, vendors could no longer offer universities significant discounts on main frames. These simultaneous events caused many university computing facilities to deteriorate.

I think the universities are partly to blame for not having taken the responsibility to keep their facilities current. Part of the problem we have today, that researchers are not able to use supercomputers, is that many previously lost interest because the facilities weren't available.

One possible condominium approach, which places responsibility on user institutions, is an approach whereby an institution would divert about 10 percent of its total computing budget toward the use of supercomputers.

You have heard from other witnesses that no single university, perhaps, can afford a supercomputer by itself. But if one takes, for example, 10 universities in the Northeast, one finds they spend roughly $70 million per annum on computing. Ten percent of that number, about $7 million, is enough to run a supercomputer center, including recapitalization, and maintenance of software.

The problem is how to get started. It is difficult to divert 10 percent of an institution's budget immediately. I propose a phased approach, which allows universities to ramp-up to this 10 percent target over 5 years.

One must simultaneously get university researchers into the supercomputing community. I think the best way, the most expedient way of doing that, will be to enhance the capabilities at existing sites, for example, NMFCC, NCAR, Los Alamos, Brookhaven, and so forth.

I would like to point out that at universities, supercomputing tends to cost about five times as much as corresponding costs researchers are charged in the national labs. Universities are either using an older main frame, where the cost per computational cycle is higher, or purchase supercomputing time from one of the few commercial vendors. Commercial costs are almost prohibitively high.

If one turns to the Japanese or to the United Kingdom or to France, one finds that university researchers have supercomputer access. In inertial fusion at the University of Rochester, about 10 percent of the annual budget is spent on a computing facility that is essentially 1970 technology.

In the United Kingdom, researchers at the Rutherford Lab, which conducts inertial fusion research, have access to a Cray. A sister organization, the Institute for Laser Engineering, in Osaka, Japan, has access to the considerable power of the University of Tokyo system which—as a Los Alamos scientist pointed out in a recent trip to Japan—usually gets the first of anything that is new, particularly if it is funded by MITI.

I believe the time has come for universities to get into the supercomputing business. If one takes a condominium approach, ultimately, funding agencies would continue to fund research, not facilities.

In a condominium approach where a university buys a share of a supercomputer facility, it is up to that university to decide how much of its portion of the facility is to be used in the training of students.

One point that Dr. Trivelpiece made yesterday is that many researchers who have grants have access. Students often do not have access. If a university owns part of a computer, it can decide what fraction of that resource it is going to use in its educational mission.

Universities like to be able to determine that for themselves.

I believe that university-industry cooperation is one of the great strengths of our country, and any national strategy we have should preserve and strengthen it. In particular, I mentioned research at my institution, is funded by both private and Federal moneys.

I sincerely believe that industry-funded university research should have an equitable treatment with that of federally funded university research. It is possible for LLE to obtain computer time from Brookhaven—or other places—for LLE's federally supported research. Although it is also possible to get computer time from these places for the industrial funded part of our research the cost must be at the "commercial rate," and the Federal laboratories can't guarantee access or availability.

I think that, across the board, we want industry-funded research and Government-funded research at universities to be able to be treated on an equal basis. I think we should have a balanced geographical nodal distribution of sites. This strategy will allow the involvement of local industry and local government to participate in supercomputer-related research.

Supercomputers aren't things like the Batavia accelerator. They don't cost nearly as much per unit. It should be possible to spread them around so people get used to them.

Why should the Department of Energy take the initiative? The reason is that the Department has a critical national role, in its nuclear weapons design and development program, in its other national security programs, and in its civilian energy programs. As a nation we need to increase our expertise in fluid dynamics, weather forecasting, astrophysics, reactor safety, reservoir modeling, weapons, advanced weapons. The DOE is the agency most likely to benefit in the near-term and requires an expertise in these areas.

I would like to conclude by saying I believe, in the last two years, particularly since the Lax report was issued, that we have done a tremendous job in this country of documenting the critical shortage if supercomputing access. I think it is time we quit documenting so that we don't continue to lose valuable time.

You mentioned yesterday, Madam Chairman, that in 2 years, may be behind other nations, particulary Japan. It may be sooner than that. I don't think we are going to catch up if we don't do something now rather than continue to document our needs.

I would estimate the cost for enhancing existing sites, at $50 million—the cost of six or seven supercomputers—to be distributed at Brookhaven, Argonne, Livermore, Los Alamos, NASA Ames, et cetera.

For new sites, I estimate the cost at $100 million over 5 years—a new site per year. The goal of this program would be to establish sites and turn over the responsibility for those sites to university consortia or a commercial contractor.

It is particularly important that university research done at either enhanced existing sites or new sites not be subject to the programmatic pressures of the host site. I think that is very important, because university programs are not able to absorb major changes in programmatic direction easily.

Sites need to be operated where these are not days where access is denied to university users, because some crisis is in progress at the host site.

Finally, let me urge, again, that the United States invest in additional research and funding to make remote access capability more effective than it is today.

Yes, satellite links exist; yes, they are expensive. We need more of these links, easily accessible, and considerably more cost-effective, to enhance access to supercomputers.

I would like to conclude at this point. Thank you.

[The prepared statement of Dr. McCrory follows:]

U.S. HOUSE OF REPRESENTATIVES
Ninety-Eighth Congress

COMMITTEE ON SCIENCE AND TECHNOLOGY

Testimony at a joint hearing before the

Subcommittee
on
Energy Development & Applications

and the

Subcommittee
on
Energy Research & Production

"Computers and Their Role in Energy Research: Current Status and Future Needs"

by

Robert L. McCrory

Director
Laboratory for Laser Energetics
University of Rochester
250 East River Road
Rochester, New York 14623

June 15, 1983

Chairmen and distinguished members of the Subcommittees, I appreciate the opportunity you have afforded me to appear before you this afternoon. In the very early days of computing, universities provided some of the most advanced computing power and support both to their students and faculties and to the rapidly developing computer industry. The situation has deteriorated significantly over the past two decades. Of the forty-two supercomputers[1] in the United States presently in use only four may be found at universities,[2] in contrast to a larger number of American-made supercomputers presently installed in foreign universities. Today, I want to address the need for state-of-the-art supercomputers for research in energy science and engineering in the United States.

My testimony will point out the urgent need for supercomputing capabilities for university research and will argue that the nation should initiate a program comprised of several parts. First, I urge that a program of limited duration be initiated to budget approximately $90 to $100 million over five years for five new supercomputing sites. I propose that the U.S. Department of Energy (DOE) take the initiative in meeting the national needs for immediate access to supercomputers by enhancing the capacity of existing sites to provide immediate capabilities, by funding additional research for enhanced remote access capability to existing facilities, and by vigorously supporting research and development activities for future computer hardware and architecture. My testimony argues that, regardless of the solution path chosen, there be an equitable treatment for industry-funded university research with that of federally-funded university research to more deeply involve the private sector in supercomputer applications and development.

I have been concerned with the need for improved access and availability of supercomputing facilities in my capacity as the director of the Laboratory for Laser Energetics (LLE) of the University of Rochester and in my capacity as chairman of the northeastern Consortium of Universities Concerned About Campus Computing (CU4C). My concern comes from both the parochial research needs of my home institution, which has national energy research responsibilities, and my involvement with leaders from many of the distinguished institutions who share this concern for the deteriorating capability and capacity of computing resources. Allow me to sketch briefly the history of my involvement in these areas before addressing your specific questions.

The University of Rochester has been involved in laser fusion research since the mid-1960's. It has pursued an expanded research program in conjunction with industry and the Department of Energy since 1972 at the Laboratory for Laser Energetics. The primary mission of the Laboratory is to determine the feasibility of inertial confinement fusion as an alternative energy source.

The Laboratory for Laser Energetics is one of the major supporting laboratories (the others include KMS Fusion, Inc. and the Naval Research Laboratory) in the National Inertial Confinement Fusion (ICF) Program. The support laboratories extend the talent and resource base of the ICF Program and provide extensive capabilities not available in the mainline programs at the nuclear weapons laboratories. The University's program is totally unclassified, and is engaged in basic research to support the programs at the weapons laboratories.

In the early 1970's, the University funded its fusion research effort almost exclusively through its industrial sponsors. As the program evolved from a bench-scale effort to a laboratory scale, we elected to meet a national goal of laser fusion feasibility research using the joint support of the Federal Government, industry, utilities, the State of New York, and the University of Rochester. Approximately $100 million have been invested in the program since 1972. Less than one-half of that total has come from the Federal Government. The highest power laser now available in the world for uniform illumination of a target is the OMEGA facility of the University of Rochester.

Federal funding for the Laboratory is derived from the Inertial Fusion Office within the National Defense Programs of the Department of Energy. Because there is no civilian inertial confinement fusion effort, the University's program, with its primary thrust in energy research, is funded from the military authorization and the Energy and Water Development appropriations accounts, in order to match DOE's organizational alignment in inertial fusion.

The Consortium of Universities Concerned About Campus Computing had its genesis in a series of discussions coordinated through Associated Universities, Inc. (AUI) which operates Brookhaven National Laboratory and the National Radio Astronomy Observatories. AUI cooperated with a number of northeast universities and other institutions in exploring the feasibility of establishing a regional computing center to meet the needs of a variety of researchers who have an urgent need for access to state-of-the-art supercomputing capability. The first of these discussions was held in March of 1982 with representatives (listed in Attachment A) from sixteen institutions.

Several conclusions resulted from that initial meeting. The participants agreed that the level of interest among the potential customers for large-scale computing, as well as potential sponsoring agencies (the Department of Defense (DOD), the National Aeronautics and Space Administration (NASA), the Department of Energy (DOE), and the National Science Foundation (NSF)), justified the formation of a more formal group to address further approaches and strategies to provide a high speed computing capability to universities. (AUI stated that it was not the proper organization to assume these responsibilities.)

Following this meeting, additional discussions were held with Mr. Jack T. Sanderson, Assistant Director for Engineering at the NSF, and other NSF staff members. Various combinations of Federal Government, industry, and university cooperative arrangements were suggested for supplying large-scale computing resources to researchers. NSF and DOD formed a workshop to document the need for large-scale computing to deal with problems in science and engineering, to discuss alternative arrangements for providing computing capacity for these problems, and to discuss the future of computer technology in the United States.

In May, 1982, representatives (listed in Attachment B) from sixteen universities and representatives from the Office of Naval Research, the NSF, the National Radio Astronomy Observatories, and Brookhaven National Laboratory, met at the Massachusetts Institute of Technology (MIT) to explore further the needs of the academic community for improved access to large-scale computing and to investigate mechanisms for increased funding for such computing at both the educational and research level. This meeting resulted in the formation of CU4C.[3] Fourteen institutions (Brown University, Carnegie-Mellon University, City University of New York, Columbia University, Cornell University, University of Maryland, Pennsylvania State University, University of Pittsburgh, Princeton University, University of Rochester, Rockefeller University, Rensselaer Polytechnic Institute, State University of New York at Buffalo, and Woods Hole Oceanographic Institute) are signators of "Computers and University Research - A White Paper" (Attachment C). The purpose of the consortium is to coordinate institutional efforts to address solutions to computing deficiencies in academic and educational high-speed computing.

At the MIT meeting, concern about the rapidity with which universities could address the pressing problem of access to supercomputing resources was central to the discussions. The attendees at this meeting cited that the need for modern computing capability and capacity was amply documented by the NSF report[4] and other reports.[5-8] Prof. Peter Lax from the Courant Institute of New York University attended the meeting and stated that his newly formed panel would be furnishing additional documentation of computing needs in science and engineering. Prof. Lax believed that the need for high-speed computing was critical, but that the greatest challenge for his panel would be suggesting timely and fiscally realistic solutions for access to state-of-the-art computing, as well as proposing levels of funding appropriate for additional future computer research and development.

A majority of the CU4C participants believed that a combination of government and university funding would be required to accelerate progress in providing computer capacity. Federal support for the establishment of shared facilities would be essential because of the heavy initial investment. Costs of creating new regional facilities to serve the national needs for access were estimated to be $90 to $100 million over five years (five centers).

Key to the establishment of successful centers, if organized as "condominiums", would be the long-term (five to eight years) institutional commitment of participating institutions for support of such centers. Independent of institutional size, a long-term commitment to divert approximately ten to fifteen percent of an institution's overall computing budget to support a shared facility was suggested as a desirable approach. The redirection of resources at most institutions, recognized as a most difficult task during this uncertain financial era, would require three to five transition years to come to full scale and would require executive action. It is important that a regional center not be initiated unless it has firm financial support ab initio, and, further, has a demonstrated demand to ensure its success. It is my understanding that there may be some center or centers already established which are undersubscribed and are in some financial difficulty. In diverting an institution's computing budget to support a shared facility, it is recognized that the transition requires government support to initiate the process, but the long-term goal of these multi-institutional regional centers would be to establish self-sustaining facilities.

An argument in favor of regionally-managed supercomputing centers is that, ultimately, the participating institutions should not be as heavily reliant on the Federal Government for support of facilities. Naturally, the institutions would continue to rely on many federal agencies to continue to fund research in a variety of disciplines, but it is believed that the agencies' responsibility would be to fund specific research proposals rather than facilities. Because the new generation supercomputers are more cost-effective, on the basis of cost per computational cycle, some savings may actually accrue to agency budgets. A properly managed center will provide for both hardware and software replacement to maintain state-of-the-art capability on a continuing basis.

The report of the panel on Large Scale Computing in Science and Engineering, sponsored by DOD and NSF in cooperation with the DOE and NASA, notes that "The United States became, and continues to be, the world leader in the use, development, and marketing of 'supercomputers', the machines that make large-scale computing possible. In the 1970's, the U.S. Government slackened its support, while other countries increased theirs."[2] Almost simultaneous with decreased government support of university centers, computer vendors reduced the attractive discounts that were previously available to universities for mainframe computers. There appears to be institutional apprehension that the assumption of prolonged government sponsorship of facilities might lead to a repetition of the events just cited. I believe that research and educational institutions must actively shoulder the responsibility for maintaining state-of-the-art computing capability, however government should have the responsibility to initiate the process. Decentralized, shared centers can be financed by the participants, and at the same time accommodate orderly capital equipment replacement and expansion.

It is especially important to realize that short-term economies and over-reliance on federal agency support may not be the best long-term strategy. Universities should be encouraged to commit their considerable resources to the development of scientific computing. In order for them to commit funds on a long-term basis, it is necessary for shared resources to be generated with a clear and participatory management structure.

The long-term vitality of the nation's scientific computing community, both at universities and at research centers, must be considered. It should be recognized that typical computational costs for large-scale problems at universities are running about five times the comparable costs at government-supported supercomputer installations. For research where government support is unavailable, universities are either using older generation mainframes, where the cost per computational cycle is actually higher than the corresponding cost on a supercomputer, or they are purchasing time from a commercial vendor. Typically, commercial rates are prohibitively expensive for very large computations. Universities in the United Kingdom, the Federal Republic of Germany, and Japan have recognized the central importance in a technologically based economy of furnishing supercomputer power to students and researchers in universities. In order for this country to maintain its international leadership position in science and engineering, its universities must be able to use cost-effective computer resources.

R. Ewald, B. Buzbee, and J. Worlton of the Los Alamos National Laboratory reported in April, 1982 on the reality of the aggressive Japanese effort to occupy the world premier position in the design and development of supercomputers. The Los Alamos scientists correctly observed that increased access to supercomputers is required to foster academic research in high-speed computing. In Japan "...the University of Tokyo consistently gets the first production model of any advanced computer whose development has been funded by MITI" (Ministry of International Technology and Industry).[9] University researchers in the United Kingdom have had access to state-of-the-art computing for a great number of years; the University of London Computing Center serves over ninety user institutions, which have remote access capability through a variety of different front-end machines. The NSF panel report gives many other examples of the strong government support of high-speed computing in other countries.[2]

In my experience as the director of the Laboratory for Laser Energetics, I have found differences in computing resources available to energy researchers in other countries with those of my institution. The Institute for Laser Engineering (ILE) in Osaka, Japan, is engaged in a large, multi-faceted effort of research in inertial fusion. The ILE computational physicists have access to the vast computing power of the Japanese university system.[9] Inertial fusion research scientists in the United Kingdom, both those at the central laser facilities housed at the Rutherford Laboratories, as well as those in universities, have access to supercomputers for their research. The National Magnetic Fusion Energy Computer Center (NMFECC) located at the Lawrence Livermore National Laboratory is accessible, in an interactive mode, to U.S. magnetic fusion researchers at many locations. Because LLE is engaged in research in inertial fusion and has industrial involvement, our research scientists do not have access to these large scale machines.

At LLE, we purchased a Control Data Corp. (CDC) Cyber 175 in 1979 in order to accommodate the prodigious computational requirements of inertial fusion research. The large fixed costs of running such a dedicated facility consume approximately ten percent of LLE's annual operating budget. We purchased the machine because our private funding for some of our research could not be used at favorable per-hour rates at existing government-owned facilities. Research connected with our federally-supported research could be accommodated at some sites, but only on a low priority, time-available basis. Presently, LLE's requirements for both computing capacity and capability have exceeded our present facilities. It is not economical for us to purchase a new machine, because we cannot take advantage of the favorable economies of scale without entering the supercomputer market. We cannot acquire a supercomputer because our university cannot afford the cost nor fully use the tremendous power the computer offers.

A supercomputer, such as the Cray Research, Inc. X-MP or the CDC Cyber-205, is prohibitively expensive for the LLE effort, but should be accessible through a cooperative arrangement with other universities or through federally-established sites, or both. The historical impediments to our obtaining the economies of scale available through larger centers, as well as our present dilemma, attest to the desirability of cost-effective access to high-speed computing which recognizes the desirability of an equitable treatment for industry-supported research with that of government-supported research at universities.

Present efforts in the applied sciences are evolving naturally to a stage where the most important problems require a numerical solution. Problem complexity arises not because of lack of judgment in simplifying the problem, but inevitably as a problem matures.

Computer modeling is relatively inexpensive compared to the costs of designing and constructing hardware for actual experiments or devices. Advances in the applied sciences are more rapidly available and cost less as modeling techniques and capabilities increase. To maintain a participatory role in state-of-the-art problems, even the best investigators must have the resources and facilities available to accomplish their tasks. The present situation is critical, particularly in the fields of fluid dynamics, numerical weather prediction, astrophysical simulations, reactor safety, reservoir modeling, and aerodynamics. Expertise and capability in a combination of all these disciplines is required to maintain a position of leadership in both civilian energy research and in the design of advanced nuclear weapons.[10] Additionally, universities have the responsibility to train students capable of using modern tools to solve problems of current scientific and engineering interests. The United States should aggressively seek to maintain its leadership position in the physical sciences and engineering by providing the best possible resources to its engineers and scientists, and ensure that they are adequately trained using state-of-the-art facilities.

NSF's Subcommittee on Computational Facilities for Theoretical Research proposed funding in the amount of $10 million per year for computational physics. The Subcommittee noted that "substantially larger funding will be needed for the long-term health of this vigorous new mode for physics research."[4] The Subcommittee's report importantly distinguished between computational capacity and capability and noted that increased capability, in particular, is required to address frontier problems. Capability is a measure of the highest degree of complexity of a single job that a computer can process in a given time. Capability cannot be increased by increasing the quantity of like computers. Enhanced capability requires both faster instruction execution speeds and increased computer central memory capacity. This distinction is particularly important to the Department of Energy in assessing future needs for computing in energy-related research. Because DOE's research is substantially broader in scope than that considered by the NSF study,[4] I estimate that current and future annual funding computer requirements is probably a factor of three to five times greater.

It is necessary to plan not just for the near-term (1984-1990), but also for 1990 and beyond. While a 100 GFlop (10^{11} floating point operations per second) machine may serve as the appropriate hardware in the future, it is necessary to have a long-term picture of how these resources should be distributed.

University-industry cooperation is one of the great strengths of our country, whose impact on present problems should not be underestimated. A national strategy should preserve and strengthen this collaboration. Mechanisms which ensure the equity of industry-supported research at universities with that of federally-supported research should be an essential ingredient of a national program. Roland Schmitt has recently observed that "There is one area, though, where government and industry have a good record of technology transfer. Supercomputers are a prime example of a technology in which defense can take the lead. They would clearly benefit the military, but at the same time they would give U.S. computer firms the opportunity to advance the state-of-the-art in ways that most of them cannot afford at this point, because even though such computers are important, the civilian market for them is not large."[11] I believe that I am joining Mr. Schmitt, CU4C, and Mr. Kerr[12] in the recommendation that government, industry, and university cooperative arrangements should be put in place now to assure the continued technical leadership of the United States in national defense and energy programs. Without the assurance of federal support for the supercomputational facilities, these cooperative arrangements will not be realized.

In the short-term, it is perhaps expedient to provide enhanced supercomputing capability at a variety of national sites (e.g., those at LANL, LLNL, the National Center for Atmospheric Research, NASA Ames Research Center, Brookhaven National Laboratory, among others). I estimate that the one-time costs associated with these enhancements is of the order of $50 million. These institutions are in the best position to alleviate the problems already upon us. Short-term solutions should not be viewed as a permanent solution to the problem of access, however. Although potent efficiency arguments can be made for continuing to support existing centers with a demonstrated experience and management record, the long-term goal of a more balanced geographical nodal distribution should be part of the national strategy. In the long run, this goal will encourage industries and local governments to devote additional resources to high-speed computing and computer development.

The diversity of university research requires that its scientific computing be conducted in an open environment, including easy and immediate on-site user access. Many of the existing sites must restrict access for national security reasons. I recommend that DOE fund five more open sites for research which do not require classification protection.

The role of university-based scientists and engineers in a national plan must be specified clearly and completely. If we as a nation wish to re-establish the partnership of industries and universities in large-scale computer development and utilization, the detailed management and resource allocation of major computer nodes must be defined in such a way as to ensure that university-based research is not subject to the programmatic pressures of host sites. The fragility of university research programs requires that they not be subjected to major swings in programmatic direction. Any national based network should provide long-term assurances of continuity of quality and quantity of service.

An effective national network should accommodate a combination of approaches for effective access, as recommended in the NSF report.[2] The report recommended enhancing the capabilities of existing sites with supercomputing capacity, providing supercomputers to selected government laboratories without such facilities and to make these new facilities available to the broader research community through networking, and the establishment of additional regional centers at selected universities. Recommendations are still urgently needed to achieve this goal, as well as to ensure an equitable resource use between nodes.

I should like to conclude by summarizing my response to your questions. There now exists substantial documentation on the critical shortage of access to supercomputers. The involvement of the academic community in research which urgently requires supercomputing capability has deteriorated. The past two years have been spent in documenting the existing needs in many disciplines. Of the federal agencies involved in these studies, the Department of Energy has perhaps the most central role, both for historical reasons and, more importantly, for its requirements for present and future needs in energy research and development. The crucial role of the Department of Energy in nuclear weapons research, its other national security responsibilities, and its civilian energy research programs, places a most awesome burden on the Department. I believe that it is time for action, and that continued studies and reports will only exacerbate an already critical situation. The United States is losing valuable ground each day we delay in closing the technological gap in supercomputer availability and computer research and development.

Since the publication of the NSF report, it has not been clear which various federal agencies should shoulder the responsibilities of implementation. I strongly recommend that the Department of Energy take the initiative and adopt the following approaches to meet the national needs.

1. Enhance the capabilities of its existing sites to provide access to a broader research community.
2. Initiate a program of approximately five years duration to budget for new condominium sites ($90 to $100 million over five years for five new sites).

3. Support research and development activities for future computer hardware and architecture.
4. Ensure, regardless of the solution path choosen, that there be an equitable treatment for industry-funded university research with that of federally-funded university research to deeply involve the private sector in supercomputer applications and development. Isolation of government, university, and industry efforts in computer research and development should be avoided. A program structured to reap the benefits of the cooperative efforts of these three components should create the momentum required to guarantee the premier position of the United States in the international community in research and engineering. (This strategy will ultimately benefit the Department of Energy.)
5. Fund research to improve remote access capability to state-of-the-art facilities for researchers in industry, government, and universities.

Thank you for the opportunity to testify before you this afternoon.

REFERENCES

1. E.W. Kozdrowicki and D.J. Theis, "Second Generation of Vector Supercomputers," Computer 13, 71-83 (November, 1980); E. Swartzlander, "Supersystems: Technology and Architecture," IEEE Trans. Compat. C-31, 399-409 (1982).
2. "Report of the Panel on Large Scale Computing in Science and Engineering," P.D. Lax, Chairman (National Science Foundation, Washington, DC; December 16, 1982).
3. J. Walsh, "Universities Seek Access to Big Number Crunchers," Science 217, 339 (July 23, 1982).
4. "Prospectus for Computational Physics," Report by the Subcommittee on Computational Facilities for Theoretical Physics to the Advisory Committee for Physics, W.H. Press, Chairman (Division of Physics, National Science Foundation, Washington, DC; March 15, 1982).
5. W.C. Rheinboldt et al., "The Need for Access to Large Computers at the University of Pittsburgh" (1982, unpublished).
6. Report of the MIT Comprehensive Planning Committee for Information Processing for the Next Decade, MIT (1979, unpublished).
7. K. Wilson, "A Push for Major Improvements in Both Large-Scale Computing Capability and Simulation Techniques for Physical Processes," Cornell University (July 15, 1981; a memo to users of large-scale computing for simulation in government, industry, and universities).
8. "Magnetic Fusion Energy and Computers: the Role of Computing in Magnetic Fusion Energy Research and Development" (Office of Fusion Energy, U.S. Department of Energy; October, 1979).

9. R. Ewald, B. Buzbee, and J. Worlton, "Foreign Travel Trip Report: Visits with Japanese Computer Manufacturers, February 1-10, 1982" (Computing Division, Los Alamos National Laboratory, Los Alamos, New Mexico; April 5, 1982, unpublished).

10. "Nuclear Weapons and Computers, Summary: the Role of Large Scale Computers in the DOE Nuclear Weapons Program," the Gold Report, prepared for the Director of Military Applications (U.S. Department of Energy; March, 1978).

11. R.W. Schmitt, "Building R&D Policy from Strength," Science **220**, 1013-16 (June 3, 1983).

12. D.M. Kerr, letter to the Honorable Melvin Price, U.S. House of Representatives (April 16, 1982; unpublished).

ATTACHMENT A

Planning and Discussion Group
AUI Headquarters, Washington, D.C.
March 19, 1982

Dr. Martin Blume
Associate Director, Low Energy Physics & Chemistry
Brookhaven National Laboratory
Upton, NY 11973

Dr. W. Robert Burns
Head, NRAO Computer Division
National Radio Astronomy Observatory
Edgemont Road
Charlottesville, VA 22901

Dr. Max Goldstein
Director, Computing Facilities
Courant Institute of Mathematical Sciences
251 Mercer Street
New York, NY 10012

Dr. Gerald Guaralnik
Physics Department
Brown University
Providence, RI 02911

Mr. Donald K. Hess
Vice President for Campus Affairs
University of Rochester
Rochester, NY 14627

Dr. Robert E. Hughes
President
Associated Universites, Inc.
1717 Massachusetts Ave., NW
Washington, DC 20036

Dr. Kenneth King
Vice Provost
Cornell University
Ithaca, NY 14853

Dr. Robert L. McCrory
Laboratory for Laser Energetics
University of Rochester
Rochester, NY 14623

Dr. Joel Moses

Building 38-4 01
Massachusetts Institute of Technology
Cambridge, MA 02139

Dr. Steven A. Orszag
Professor of Applied Mathematics, 2-347
Massachusetts Institute of Technology
Cambridge, MA 02139

Dr. Alan Perlis
Department of Computer Science, Durham Laboratories
Yale University
New Haven, CT 06520

Dr. Anthony Ralston
Department of Computer Science
State University of New York
4226 Ridge Lea Road
Amherst, NY 14226

Dr. George B. Rybicki
Associate Director, Theoretical Astrophysics
Harvard-Smithsonian Center for Astrophysics
60 Garden Street
Cambridge, MA 02138

Dr. Paul Schneck
Code 500
Goodard Space Flight Center
Greenbelt, MD 20771

Dr. Thomas D. Taylor
Applied Physics Laboratory
The Johns Hopkins University
Johns Hopkins Road
Laurel, MD 20707

Dr. Conrad A. Wogrin
Professor of Computer & Information Science
University of Massachusetts
Amherst, MA 01002

ATTACHMENT B

ATTENDEES

MIT Meeting
May 3, 1982

Brookhaven National Laboratory

Dr. R. Peierls (Head, Applied Mathematics Division)

Brown University

Dr. M. Glicksman (Provost)
Professor G. Guralnik (Department of Physics)

Carnegie-Mellon

Dr. D. Van Houweling (Vice Provost Computing)
Professor G. Fix (Head of Mathematics Dept.)

City University of New York

Dr. I. Fuchs (Vice Chancellor of University Systems)

Columbia University

Dr. B. Gilchrist (Director of Computing)

Cornell

Dr. K. King (Vice Provost for Computing)

Courant Institute of Mathematical Sciences

Professor O. McBryan
Professor P. Lax

Harvard University

Dr. R. Leahy (Associate Dean)
Dr. G. Rybicki (Astronomy Department)

Johns Hopkins Univeristy

Dr. M. Karweit (Director, Engineering Computer Facility)
Dr. T. Taylor (Applied Physics Laboratory)

University of Massachusetts

Dr. C. Wogrin (Director, University Computing Center)

MIT

Professor F. J. Corbato (Electrical Engineering & Computer Science)
Professor J. Moses (Head, Electrical Engineering & Computer Science)
Professor S. A. Orszag (Applied Mathematics)

National Radio Astronomy Observatory

Dr. W. R. Burns (Head, Computer Division)

Princeton University

Dr. J. Poage (Computer Center Director)

Rensselaer Polytechnic Institute

Mr. J. Moss (Director, Office of Computer Services)

University of Rochester

Dr. R. L. McCrory (Laboratory for Laser Energetics)
Dr. D. Stonehill (Vice Provost for Computing)

Rockefeller University

Dr. M. Ferentz (Director of Computing Services)

Woods Hole Oceanographic Institute

Dr. W. S. Little, Jr. (Manager, Information Processing Center)

Yale University

Professor A. Ferguson (Associate Provost)

Fred C. Hart Associates

Dr. B. Miller (Vice President, Energy Programs)

National Science Foundation

Dr. L. Lee (Program Coordinator Computing Facility NCAR)
Dr. E. F. Infante (Division Director - Math. Sciences)

Office of Naval Research

Dr. J. C. T. Poole (Director, Mathematical and Information Sciences)

ATTACHMENT C

Computers and University Research

A White Paper

June, 1982

High-speed digital computation is a mainstay of modern scientific investigation. It is now commonplace to find research that demands access to computing power that only twenty years ago was not available at any cost. In the context of university research, faculty and students routinely rely on the availability of such systems to complete their work.

In the past, the university community's use of computers tracked the growth of computing power. In many cases, the sophistication that university investigators brought to the use of existing systems was the impetus for development of new technology. In fact, there is a prima facie case to be made that the growth of computer technology has been driven as much by the demands of university researchers as by any other factor. That now appears to be changing.

The cost of the latest generation of computers represents a risk that most universities will not take in today's fiscal environment; and the alternative to buying a system, i.e., the purchase of services on a reimbursable basis from commercial vendors, is often prohibitively expensive. This dual constraint has had a disturbing impact on the university research that relies on advanced computational capability.

Many investigators who depend on access to state-of-the-art systems are relinquishing their leadership position as innovators in the use of computers to advance research. The implications of such a shift cannot be ignored. Nor should they be underestimated. As university investigators turn away from any class of problems because they lack the tools to attack them, all of science suffers. Students are shortchanged in their training because of limited exposure to advanced technology and science is shortchanged because students and faculty are denied the opportunity to address challenging problems.

In recognition of this emerging trend, the signators to this document have agreed to coordinate their efforts in bringing the problem to the attention of the established organs of government and science through a joint enterprise known as the Consortium of Universities Concerned About Campus Computing (CU4C). The purpose of the consortium is:

o to provide a focal point for assessing the status of campus computing,

o to relate that status on a continuing basis to the needs of the university research community,

o to present findings in the appropriate forums of government and science, and

o to propose and suggest suitable mechanisms for addressing deficiencies that are identified.

Mrs. Lloyd. Thank you very much. Your testimony is excellent. Thank you.

I am going to pose a question for all three of you to think about. Yesterday, Dr. Trivelpiece suggested a computer science program for the next supercomputer should be focused on a set of specific scientific problems of national importance.

Now, such an approach would necessarily force cooperation between university labs and the Government, and might develop more than one kind of supercomputer. How would you react to this approach?

Dr. Fernbach. I would like to say that a supercomputer is a supercomputer is a supercomputer. I feel that we don't have to design a computer for a specific problem. A supercomputer, if it is well designed, will do most of those problems that we think are critical.

The problems don't differ that much. I believe that the two most important jobs that we are faced with are in the NASA program, and in the DOE program for the nuclear weapons and reactor development.

The problems are very similar. I believe that if we can get the supercomputers we need, they will be able to solve those problems.

Mrs. Lloyd. Thank you.

Dr. Olson?

Dr. Olson. Oh, I think what we are talking about here are number-crunching computers, and not symbolic computers. And I think if we are talking about a very fast numerical processor, the key to what discipline it is applied to is how easily is that discipline adapted to numerical analysis?

Once you have done that, then the supercomputer is just the machine to more rapidly analyze your data or bring it to some reasonably well-defined result. Our own supercomputer facility is primarily used by engineers and physicists, two of the people of our chemistry department, but 2 weeks ago, we had a group come over from our Life Science Group asking if it was possible to numerically represent an endocrine system in a mammal.

Well, that is a big job for the mathematician, but if they could do that, they would then be able to model on a Cyber 205 a mathematical representation of a living system and do all kinds of experiments on it, on the computer, just the same way that the cloud physicist models his cloud formation problems today.

So I agree with Dr. Fernbach, a supercomputer is a supercomputer.

Mrs. Lloyd. Dr. McCrory, do you think a supercomputer is a supercomputer?

Dr. McCrory. I would say something a little different. I like Al Trivelpiece's suggestion, even if it were never carried out. One should identify the critical national problems one wants to solve. What are the urgent problems of our time? It is then proper to ask to what extent present hardware and software effectively address those problems, or whether we need something different.

Is the vector architecture really the one we need most for what are the most urgent scientific problems or do we need a mix of parallel processors and vector processors to address key scientific and engineering questions? We won't know until we decide which problems are the most important. Are we making the solution fit the problem, or the problem fit the solution?

Mrs. **Lloyd**. Thank you.

I must move on. The next is a two-part question.

Do you think that our present educational system is producing enough adequately-trained scientists and engineers to use and support the future needs of our computer design; and do you think that limited access to supercomputers is restricting our ability to train?

Dr. **Fernbach**. I believe the latter; that the lack of access is a severe restriction. Our people, our students have no way of learning what a supercomputer is all about; how to use it, and what algorithms are needed; what software is needed; and how it could better perform their problems, their research problems.

Mrs. **Lloyd**. Well, let me go a little bit farther. Do you think they are being adequately trained up to the point where they could use a supercomputer? Are we lagging in that respect as well?

Dr. **Fernbach**. Well, I think we are lagging there, too, but that is a different type of thing. I feel that our computer science curriculum needs some revitalization. But that is a another matter, I don't believe that we are doing enough good teaching in computer science and in the use of computers.

Mrs. **Lloyd**. You know, we might have a supercomputer at every university, but that does not mean that it could be certainly used well or to its full advantage.

Dr. **Fernbach**. I think most of our students that are coming out of the universities, even Ph. D. students, don't even know what a supercomputer is. I think that is terrible.

Mrs. **Lloyd**. That is sort of the objective of my question. Dr. Olson?

Dr. **Olson**. I can give you an example of that. We are not putting out near the number of graduate students trained in writing large vector codes. For example, the only discipline on our campus where graduate students write large vector codes that can be run on a supercomputer is in atmospheric science.

Last year, the top Ph. D. student in atmospheric science was hired by Mobil Oil Co., not because he knew anything about atmospheric science, but because he could write large codes, which was just a tool that he used to get his discipline degree. But he is now working for Mobil in developing codes of reservoir simulations.

So they had to go completely out of a discipline even remotely related to petroleum engineering to get someone that could do the kind of computational work that they wanted. They literally bought him. He couldn't turn down the offer.

Mrs. **Lloyd**. Well, it is really not the objective of these hearings to get into this discussion, but I think it is very important, if you would like to elaborate further, I appreciate hearing from all of you.

Dr. McCrory?

Dr. **McCrory**. I agree with both of the gentlemen. We are not training enough students. Things cost more and more these days in a highly technologically oriented society. Modeling is considerably cheaper than cutting iron. Training students to support our needs requires improved access. The Department of Defense may not need big computers, but they need a significant numbers of people to enter careers in the defense sciences. These people need to be adequately trained using what is now known as a supercomputer, and these computers will be what DOD will be buying in significant numbers in the future.

Mrs. LLOYD. Well, if you had complete access to a supercomputer network, how would you best use it, in what areas do you think it would enhance the programs most advantageously, from the DOE perspective?

Dr. FERNBACH. Well, certainly, in the sciences, in physics, as a matter of fact, if you listen to the Nobel prize winner of last winter, he started to use computers a number of years ago, and they are beginning to use these computers very extensively in theoretical physics for the first time.

But, in other physics research, I think—in chemistry, in oceanography, meteorology—most of the sciences are very important areas for further development of students using supercomputers.

But I go further than that. I think even in economics and disciplines like that, it is going to become important and we are just on the verge of using this thing in beogenetic engineering, these people don't even know what a computer is, let alone a supercomputer, and they are beginning to start thinking in terms of very large problems.

So, I think it is going to spread wildly from the physical sciences.

Mrs. LLOYD. Dr. Olson?

Dr. OLSON. Yes, I can't imagine a discipline in the physical sciences that would not be enhanced with access and funding to a supercomputer. Just the—any problem in fluid dynamics that a supercomputer capability let's you look at three-dimensional problems, where in the past we only looked at two-dimensional problems.

In the practical applied sense, our river mechanics people are hoping to model for the first time an undeveloped river, the Niger River in Africa, and do everything that could be done to this river in the next 100 years. Power plants, diversion for irrigation, transportation, building docks.

And you can do that and simulate it and study any event on that river from a drought to a 500-year flood, and do it on a computer. It is just a good engineering example of where a supercomputer can be used.

Mrs. LLOYD. Dr. McCrory?

Dr. McCRORY. Yes, I think problems have become more complex, not because of a lack of taste in simplifying problems, but inevitably as fields mature. As one example, not only can we now do more advanced problems in fluid mechanics, but moleling reacting fluids is very important. Only because supercomputers are available is it conceivable to calculate the solution of the fluid equations in the presence of chemical reactions heat transfer, etc.

In many DOE programs, fluid flow (e.g. rock flow which is important to predict earthquakes, or a turbulent fluid flow in a heat exchanger), is important to model. Modeling many critical areas is now possible, for the first time, because of supercomputers.

Across almost the entire scope of DOE's responsibilities, I find that there are few areas which wouldn't benefit from supercomputers.

Mrs. LLOYD. Thank you very much.

Mr. Lewis?

Mr. LEWIS. Thank you, Madam Chairman.

Dr. Fernbach, you state that if left to their own devices, U.S. manufacturers of supercomputers would not move very quickly into new, improved designs. Two of the three manufacturers yesterday advised this committee that they could provide needed capabilities, that the goals were established by the user community.

Would you like to comment?

Dr. FERNBACH. Yes. They are trying to provide a capability which is fairly easy to achieve. I am speaking about a jump ahead, a leap ahead. I think the Japanese are doing that. In other words, they are going up to gigaflops capabilities, where our people talk about megaflops capabilities.

I feel that if we push them, and only if we push them into the next range of capability will they go there. It is very difficult for them to make that move on their own. They are not willing to risk it. I feel that it is important for us to subsidize part of that development.

Mr. LEWIS. I see. Dr. Olson, what amount of time is available on these supercomputers which are presently used in the university environment?

Dr. OLSON. Oh, I would say almost—three that are at the university campuses, excluding the one from NCAR, probably 75 percent of the time on each of them is available.

Mr. LEWIS. Seventy-five?

Dr. OLSON. Yes.

Mr. LEWIS. I see. OK. Thank you, sir.

Dr. Olson, your condominium proposal really is interesting. I do hope that you study it very carefully, if you have been involved with condominiums like elected officials have, you would find it may not be very practical. Make sure the rules are all right upfront.

Dr. OLSON. It doesn't hurt to get advice both ways.

Mr. LEWIS. Dr. Olson, do you know whether 75 percent of that time has been made available to the Department of Energy, which is a user community?

Dr. OLSON. Well, we have met with some of the national lab personnel that do not have class 6 machines, but there are budget problems and problems of a computer center director at a national lab taking a part of his probably already-strapped budget to buy time outside of his facility isn't very feasible.

I think what we are talking about, to provide access to either the university community or the national lab scientists is an additional availability of funds.

Mr. LEWIS. OK.

Dr. McCrory, a comment was made, or has been made, that the university community has not distinguished itself in management of computer facilities. Would you like to comment on that?

Dr. MCCRORY. I think part of the answer to the question is: Is it a question of the chicken or the egg? An argument made by managers of existing large sites, once they have trained people that are capable of running such a large computer facility, is: "See how well they do it"?

I don't think they always did it well, and in particular, most of what has been suggested today for new sites are shared facilities. Ab initio, those proposed centers will have to have an effective management plan.

If you placed a Cyber 205 or a Cray-1, into the hands of my university or Harvard or MIT, I think they would come up with a way of running and managing the facility that would please most critics in a short amount of time. I don't think that is a problem.

Mr. LEWIS. Thank you.

Dr. Olson?

Dr. OLSON. I would like to comment on that question. Part of the problem of the disintegration of central computing facilities at universities is the funding agencies themselves, which have made the acquisition of VAX computers, minicomputers, available to individual, principal investigators, and scientists on the campus, and effectively drained away support from a strong, central computing center.

I think our most prestigious universities have suffered the worst from this, and that there are not large, strong central computing centers on the 10 largest research universities in this country, but there is a great proliferation of VAX's and small minis in the individual labs.

When we criticize our university computer center operators, fighting a situation where the organization for whom they work continues to drain away their customers.

Mr. LEWIS. I see. Dr. Fernbach, would you care to comment?

Dr. FERNBACH. Well, in the first place, I don't think a university ought to be in the computer business. It should not be selling service. The university is a research or educational institution. If it needs computer service, it ought to be getting it outside. I really do not believe any of these schools will ever run a computer center effectively, especially Harvard.

Mrs. LLOYD. I think we are going to have to vacate this room in about 5 minutes, so I do want to thank you for being here today, and your participation. We will hold the record open for an additional week for other comments to be included in this hearing.

We will be submitting questions to you in writing for also inclusion in the record.

Thank you. Thank all of you for being here and your participation. Without further comments, the subcommittee is adjourned.

[Whereupon, at 3:55 p.m., the subcommittee was adjourned.]

APPENDIX

Additional Material for the Record

NATIONAL SCIENCE FOUNDATION
WASHINGTON, D.C. 20550

June 23, 1983

OFFICE OF THE
ASSISTANT DIRECTOR
FOR MATHEMATICAL AND
PHYSICAL SCIENCES

Honorable Don Fuqua
Chairman, Subcommittee on
Energy Development and Applications
House of Representatives
Washington, D.C. 20515

Dear Mr. Fuqua:

Dr. Knapp has asked me to reply to your letter of June 6, 1983, in which you invite testimony from the National Science Foundation on "Computers an their Role in Energy Research: Current Status and Future Needs" to the joint hearings of the Subcommittees on Energy Development and Applications and on Energy Research and Production. The Foundation is grateful for thi opportunity to comment on a topic of significant importance for science, engineering, and advanced technology.

The availability of sophisticated computational facilities during the past thirty years has been crucial to the Nation's advances in science, engineering, national security, and industrial productivity. The weapons program of the Department of Energy, cryptographic analysis, and weather forecasting are essentially dependent on computers. The use of computational methods in the aerospace, petroleum, semiconductor, and nuclear industries has contributed substantially to advances in these areas. In basic research, scientific computers are as important to the theoretician as sophisticated instrumentation is to the experimental scientist. Moreover, advanced computational facilities allow for simulation of complicated scientific and engineering phenomena, not tractable by either theoretical or experimental methods.

It follows that the availability of computational facilities is crucial to the mission of a number of Federal agencies. This is particularly the cas for the Department of Energy, with its programs in weapons development and magnetic fusion research; it is also the case for the Foundation's researc programs in the atmospheric sciences and for the NASA programs in aerodynamics. The importance of modern sophisticated computational facilities has also become increasingly clear for all programs of basic research, especially in the mathematical, physical, and engineering sciences. Both the Department of Energy and the Foundation have large programs of research in these areas.

In 1982, the widespread and increasing need for computational facilities t carry out basic research programs led a number of Federal agencies, notabl the Department of Energy, NASA, and the Department of Defense, to join the National Science Foundation in conducting a study of this problem. The result of this study was the Report of the Panel on Large Scale Computing in Science and Engineering, commonly known as the Lax Report. This study, and its supporting documents, clearly indicates that significant scientific opportunities cannot be pursued without the needed computational facilities. It is also clear that the lack of investment of sufficient funds by the Federal Government, during the past decade, in scientific computational facilities (and in all modern scientific instrumentation) must be redressed if the U. S. is to retain the scientific leadership it has enjoyed in the past forty years.

Within this context, let me attempt to reply to the questions posed in your letter.

1. It is quite clear that the number of state-of-the-art supercomputers made available to energy science researchers must be significantly increased. A recent study by DOE has shown that the needs of researchers supported by the Basic Energy Sciences programs require at least one additional Class VI supercomputer in FY 1984. For FY 1985, another similar facility will, in all likelihood, also be needed. These are requirements for the Basic Energy Sciences programs only; increasing needs are also evident in the other large research and development programs of DOE, which are pressing the capacity of the many supercomputers available to them.

   Experience at the National Science Foundation is similar. During FY 1983, we have doubled to two the number of Class VI supercomputers for atmospheric sciences research at the NCAR (National Center for Atmospheric Research) facility. In an extremely short period of time, less than three months, both of these supercomputers are at full capacity. Neither of these machines is available to researchers supported by the Foundation's other programs that correspond to the Basic Energy Sciences Program at DOE; it is estimated that these researchers have an immediate need for at least two supercomputers.

   At both DOE and NSF, it is also quite clear that there are scientific requirements for all types of computational facilites, not only for supercomputers; and that these computational needs must be balanced against requirements for other types of modern instrumentation.

2. Ever since the early days of the Atomic Energy Commission, the Department of Energy has played a very major role as a buyer and developer of computers for applications to scientific and defense problems. It is to be hoped that this role will continue, especially because of the computational talents and experience that reside at the National Laboratories. In scientific computing, the National Laboratories have played a major, almost unique, part in the development of the operating systems of the CRAY computers and in developing software and algorithms for their appropriate use. The

current distribution of supercomputers in the U. S., with its preponderance at the National Laboratories of DOE, also indicates that the Department is a major market for supercomputer manufacturers. As such, it exercises considerable influence on the R&D activities of the private manufacturers. DOE has consistently been a leader in the computational sciences; its mission and needs indicate that this leadership role should continue.

3. The current level of support for computer R&D is at an all time high in both federally funded programs and in the private industrial sector. However, critical needs, such as basic research for numerical computing and the training of necessary research personnel, are not sufficiently emphasized. The domestic computer industry must allocate its scarce research and development funds to meet all commercial opportunities and competitive challenges, and these are numerous. Research and development for numerical computing and supercomputers enjoys a priority which reflects competitive conditions, and does not and could not reasonably be expected to reflect the entire national scientific and security interest. Similarly, the computer research programs sponsored by the Department of Defense, most notably by the Defense Advance Research Projects Agency, emphasize architecture, software, and algorithms based on artificial intelligence techniques; defense requirements dictate de-emphasizing numerical techniques.

   Basic research for numerical computers and numerical computing, essential to scientific and engineering research, is supported by the programs of DOE, NSF and, to a much smaller extent, by NASA. This is an area of significant scientific opportunity, one in which there is a severe shortage of appropriately trained personnel, owing to heavy demands by industry, private corporations, National Laboratories and universities. In Budget Requests to the Congress for FY 1984, both DOE and NSF have emphasized research on numerical computing and computers, and the need for increased access to computational facilities by research scientists and engineers. It is our belief that these requests for FY 1984 must be followed by significant continued expansions in programs of research and of access to supercomputers.

4. The private sector is developing the type of computers needed for forefront research in DOE and NSF programs. Control Data Corporation, Cray Research, and Denelcor completely dominate the national and international supercomputer markets. However, Japanese manufacturers, aided by significant research and development funds from the Japanese Ministry of International Trade and Industry, are poised to enter these markets with machines that appear to be comparable to the latest American models. Moreover, the Japanese manufacturers are engaged in a cooperative R&D effort to develop by 1989 a supercomputer 1,000 times more powerful than the CRAY-1S computer. No comparable American programs exist; neither is it clear that they should exist. However, knowledgeable analysts are in general agreement that the present level of American R&D on supercomputers in private industry is not adequate

> to meet the anticipated scientific and commercial challenges. Therefore, they strongly advise agencies such as DOE and NSF to increase their funding of relevant basic nonproprietary research at the universities and at the National Laboratories.

In FY 1984, the National Science Foundation will expand its programs for computer research, computational mathematics, and access to scientific computing equipment for the broad spectrum of scientists and engineers whose research it supports. The Foundation has also established a Working Group on Computers for Research to define a specific, detailed program of future action to meet the computational requirements of science and engineering research, to appropriately support basic research essential to future generations of computers, and to train sufficient numbers of scientific personnel in the computer and computational sciences. The task of this Working Group, which has developed close ties to similar groups in DOE, NASA and DOD, will be complete in July. However, its early findings mirror those presented at your joint hearings.

- o The limited computational facilities, both in capacity and capability, available to university researchers have hampered their scientific work, resulting in significant research opportunities being missed; problems of scientific, engineering, and technological importance are not being addressed because of the inadequacy of access to state-of-the-art computers.

- o The inadequate access to computational facilities for university researchers not only discourages computer and computational research, but also results in the lack of development of appropriately trained scientific personnel, much in demand in industry and at the National Laboratories. The future well-being of U.S. science and engineering depends on the development of a large, imaginative, and skilled community in the scientific use and design of computers.

- o Significant scientific and engineering opportunities will be made possible by computers much larger and faster than those available today, and by appropriate mathematical, algorithmic, and software tools. It is therefore necessary to substantially increase basic research in computational mathematics, software, and algorithms, and in other areas of research fundamental to the design and implementation of future supercomputer systems.

It is the intention of the Foundation, in coordination with the other Federal agencies and DOE in particular, to provide substantially increased facilities for large scale computing and for networking in the academic science research communities, and to actively pursue a significantly increased program in computer and computational science. This will have to be done within the context of the total resources available, balancing the needs and opportunities in these areas with those of all other areas of research. The detailed plans will be worked out shortly, and I will be glad to keep you informed on this subject.

On behalf of the Foundation, I thank you again for this opportunity to express our views on this important topic.

Sincerely yours,

Marcel Bardon

Marcel Bardon
Acting Assistant Director for
Mathematical and Physical Sciences

8100 34th Avenue South
Mailing Address/Box 0
Minneapolis, Minnesota 55440

Robert M. Price
President

CONTROL DATA
CORPORATION

June 15, 1983

Honorable Don Fuqua
Chairman of the Science & Technology Committee
United States House of Representatives
2269 Rayburn Office Building
Washington, DC 20515

Dear Congressman Fuqua:

Members of your staff extended an invitation to W.C. Norris to appear before your energy development applications subcommittee on June 14, 1983, to discuss the topic of computers and their role in energy research. Unfortunately, his schedule did not permit him to participate at the hearing.

In lieu of his appearance at this time, we have prepared the attached document that will communicate our position on this important topic. I request that it be included in the hearing.

Control Data is committed to support your efforts in bringing the needs of energy research to the attention of public policy makers.

If we can be of further assistance, please contact us through Lois D. Rice (202-789-6527), our Senior Vice President for Government Affairs in Washington D.C.

Sincerely,

CONTROL DATA CORPORATION

R. M. Price

Attachment (1)

cc: L. D. Rice

CONTROL DATA CORPORATION
Statement for the Hearing Record
of the Subcommittee on Energy Development and Applications

Control Data Corporation is pleased to provide a written statement for the hearing record concerning computers and their role in energy research.

Control Data has over 20 years of continuous service in providing users of engineering and scientific applications with the most powerful computers possible. In fact, Control Data developed the first computer described as having "supercomputer" performance.

Our history also includes one of consistent cooperation with and service to the federal government, to meet the needs in computing and in particular superscale computing. As part of that involvement, Control Data has for some time been concerned about our international leadership in supercomputing. As early as 1978, Control Data executives gave a series of presentations to government officials on the future of supercomputers, pointing out the need for the government to invest substantially more in the research and development process for high performance technology, as well as changing existing procurement practices, if the nation were to remain the leader in scientific computation through the next decade.

Accordingly, CDC strongly concurs with the basic premises of the Hearing Charter narrative and description supporting its purposes. In particular, we support the premises that:

- Today's supercomputers utilize vector processing. It takes resources and time to develop needed algorithms and methods for these relatively new architectures.

- Development of significant improvements for scientific users will be slow in comparison to those for the general commercial user unless the federal government becomes a major partner in new computer development.

The opinion in the Lax panel report of December 26, 1982, that "under current conditions there is little likelihood that the U.S. will lead in the development and application of this new generation of machines" is shared by many including ourselves. Our national leadership has depended and is depending upon the perservance of a handful of dedicated technologists and their supporting corporations. The demands of today's technology and that of the future, coupled with the growing international competition evidenced in commitments by the United Kingdom, Germany, France, and Japan, make it imperative that new and different approaches be found to ensure continued United States leadership.

Our statement is structured to respond to the four questions raised in the Subcommittee's Hearing Charter. It is consistent with and relates to recent presentations or transmittals we have made relative to the National Science Foundation's role in supercomputing and to DOE, DOD and NASA planning.

Control Data is pleased that the Congress and the House Science and Technology Committee, in particular, are addressing our nation's current status and future needs in such a vital area. We are available to aid this subcommittee in any way we can.

QUESTION 1: To what extent and in what quantities should state-of-the-art supercomputers be made widely available to energy science and engineering researchers at national laboratories and universities?

Control Data agrees with the Lax panel report that there is a substantive need to provide increased access for researchers to develop experience and applications breadth on supercomputers and vector technology. These needs exist in both the national laboratories and the leading universities. Programs and facilities must be provided to both government laboratories and universities. This will develop the overall expertise essential to maintaining leadership in supercomputing applications. This approach is necessary because of the great interdependence among universities, government, and industry in the development and application of supercomputers.

We believe it is necessary to provide universities with the means to operate and use supercomputers for applications research and instruction. To do this Control Data has encouraged the National Science Foundation to initiate a program for supercomputer applications research. A copy of our presentation to them is included as an attachment. This program should be to provide access to supercomputers currently installed at universities (Colorado State University and Purdue University each with CDC Cyber 205's and the University of Minnesota with a Cray 1) for which no explicit NSF funding is now available.

In summary, we believe that fiscal 1984 funds should be provided to:

- Increase access for researchers to existing university and national laboratory supercomputers.
- Increase the number of supercomputers installed in university and laboratory environments.
- Interconnect these facilities through a shared network.
- Increase supercomputer applications research and development.

QUESTION 2: What role can the Department of Energy play as a buyer and/or developer of technology for the next generation of computers?

Since the late 1970's, government procurements for advanced computers have been controlled by a "try-before-buy" policy. This policy is common in government procurements and superficially seems equitable in protecting its interest. However, it places the total risk for developing new supercomputer technology on the manufacturers without any assurance of a market. In addition, since the market for supercomputers is relatively small, the major manufacturers are increasingly unable to make the necessary investments to achieve the best performance in the time frame needed. In our opinion, the need for cooperative programs between the government and the private sector is clear. The private sector cannot continue to take the risk alone.

We believe the DOE should stimulate and facilitate development of the next generation of supercomputers by assuming the role of a predictable and dependable buyer. To accomplish this, we recommend DOE take the following actions:

- Establish the initial market for the next generation of supercomputers by committing to a substantial number of systems from the manufacturers that are able to verify their designs to meet the DOE's threshold requirements and adhere to appropriate milestones.

- Invest in joint development projects with universities and private industry for research into mathematical algorithms, investigations into common operating system software, and development of new applications using advanced vector and parallel architectures.

- Acquire more supercomputer capacity as soon as possible for both energy research and development of critical applications. Included would be the conversion of existing "lab-supported software" to operate on our supercomputers. We believe this effort would accelerate applications development for the advanced supercomputer systems.

- Provide direct funding for advanced technology research. This would help support multiple research programs to ensure the best technological approach is used to develop the next generation supercomputer.

QUESTION 3: What is the current status and future prospects of federally and privately funded supercomputer research and development?

The long-range financial health of a computer manufacturer demands development of a solid, expanding commercial user base in addition to their government business. The commercial market primarily requires turn key systems for their applications. Performance is important but secondary to their immediate production application.

By contrast, government lab priorities require high performance computational engines. By their very nature, the labs are more experimental than commercial enterprises. The labs have been willing to and often must be totally responsible for the applications software as well as the operating system software. This trade-off has allowed them to apply maximum computer performance to their specific applications.

Because of the different requirements of the commercial and government markets, computer manufacturers' development schedules are typically not in step with the federal requirements. Specifically, Control Data's supercomputer development program does not meet the government's schedule.

As stated previously, to shorten the development time, the government must provide an up-front commitment for systems and applications development. The government may also provide funds directly for technology development.

QUESTION 4: Is the private sector's research directed toward the development of large memory, high-speed calculational machines, which are specifically needed for the forefront research being pursued in DOE programs?

Control Data high technology development efforts are dedicated toward providing larger memory devices and higher speed processors for the large-scale scientific and engineering user.

As stated, however, in our response to Question 3, the timing of our efforts relative to the requirements of the government is dependent on the implementation of the recommendations in that response.

ACADEMIC SUPERCOMPUTER ACCESS: A CALL FOR COMMITMENT

APRIL 12

The National Science Foundation

by E. R. Krueger
Control Data Corporation

Thank you for this opportunity to speak to the Foundation. It's a personal pleasure for me, after so many years in higher education and some direct contact with several of you.

As many of you know, Control Data has a 25-year history of providing systems and services with an early involvement with institutions of higher education. We, accordingly, have a sincere and substantial interest in that market and its well-being. It is because of that interest and the growing national concern about super-scale computing that we are here this afternoon.

We understand that the Foundation is considering various approaches to support academic access to supercomputers. We offer suggestions on that topic. I hope our two approaches, yours and ours, will be very similar. If so, we have a good basis for further discussions as to how we may cooperate in the future.

Our message this afternoon is a simple one.

> We believe the Foundation should champion a national program for university research and development in super-scale computing applications. The program should initially provide access for academic researchers to

existing supercomputer facilities, both in universities and in government. It should include plans for additional university supercomputer sites.

Interpret this as a recommendation, a challenge, or, more appropriately, a call for commitment.

The recommended NSF program should, in our opinion, have several readily identifiable attributes. These are:

1. Build on experience and the existing base of proven capabilities.
2. Identify, as a national priority, supercomputer access and applications research.
3. Start quickly with initial focus on access to existing sites/networks for key disciplines.
4. Develop and implement a long term plan for a national, multiple university supercomputer network.
5. Foster cooperation between government, university and industry through this process.

To develop the rationale for these attributes and to lead into follow on discussions, my talk has four major elements:

1. Evidence of need for NSF and university involvement in supercomputer applications research.
2. NSF successes and strengths.
3. Suggestions for NSF program attributes.
4. Implementation candidates.

The evidence of need for supercomputers has been well documented and, based on the Lax report and ensuing activities, is well accepted. Many industrial applications require computer speeds at the limits of technology. Many of today's problems requiring super-scale computing are old friends with increased complexity or are making the transition from 2 to 3 dimensions - aerodynamic design and modeling, energy resource recovery and modeling, nuclear energy, weather research and structural design and analysis. Relative newcomers include:

1. Electronic design and automation.
2. Robotics.
3. Expert systems.
4. Image processing.

The current accepted approach to address these problems is to develop architectures incorporating characteristics of parallel, vector, and multiprocessor approaches. All of these require significant additions in software. New and modified systems and new and modified languages will have a great impact on applications software.

Our tendency nationally is to focus on the hardware and anticipate that the software will be developed by the users as the needs arise.

To date, industry and government labs have assumed the leadership in the development and application of the latest generation, class VI machines. At first glance, with over 60 machines installed including four in universities or supporting universities in the United States, the industry and its user community would appear to be doing well at advancing the state-of-the-art. The Lax report says, however, "not so", that we are in danger of losing our national leadership in supercomputing. The report cites problems in access to supercomputer resources and in the development of new supercomputers. It recommends more investment is required for increased access for scientific and engineering research; for increased research in computational mathematics, software and algorithms; for training of personnel in scientific and engineering computing; and finally for research and development basic to the design and implementation of new supercomputer systems.

Why is our combined effort insufficient to maintain our national leadership? Although industry and government labs have done a yeoman-like job in development and application of the latest generation of machines, historically computer vendors have not been primary sources of applications. We vendors have our hands full dealing with hardware, operating systems and languages.

The user community is much better equipped to develop applications and has performed well. For example, this approach has obviously worked in the rapid evolution of the minicomputer and microcomputer industries. It cannot, however, work satisfactorily in the now-emerging world of super-scale computing, a world of different architectures and large capital commitment. In addition, with the competition we face as a nation, there is not the luxury of time.

For Class VI machines, government labs are the current primary source of applications. However, their manpower is restricted to specific projects and they have a narrow applications base. Often their results cannot be generalized and supported or even shared outside.

Universities have some access to the latest Class VI equipment but only for a few disciplines. Universities recognize the need for more access but in most cases lack the resources to proceed. The Foundation does not have a program in place to assist in meeting that need.

Lost are the significant benefits to the nation which should come from substantial university involvement both as users and operators of this equipment. The expected benefits are well articulated in Professor Ken Wilson's paper included in the Lax report. In summary they are:

- System verification and enhancement
- Development of a broad applications base through instructional and research use
- Focused training of the future user base

I cannot emphasize enough the need for both supercomputer research and training. Let me quote Professors Keller and Rice when they said, "one does not devise new methods or learn to think in parallel overnight." Parallel processing is just one of the many challenges we face.

The Lax report and supporting papers give substantial credence to the great interdependence between universities, government labs and industry. The supercomputer industry looks to the mission-oriented agencies for support as customers in the processes of design and development. We look to the universities for broad applications development through research and for training through the educational process. Because of the significant capital funding requirements, the universities look to NSF for assistance. We feel this interdependence between universities, government and industry is so great in the application and development of supercomputers that reduced or limited performance by any one severely impedes the overall progress of the group. Our national leadership in supercomputing can be likened to a three-legged stool. If any one leg is reduced, the overall function is degraded. Too much reduction, the overall function is in jeopardy. We are in the latter environment today. The universities are not able to maintain their historic role in the research and development process. Government and industry cannot supplant that shortage.

We are fortunate to have two supercomputer suppliers in the United States. At this time, the only two in the world. It is ironic that more universities outside the United States have supercomputers installed than in the United States. To our knowledge, those external to the United States have without exception been partially funded by their federal governments. It appears that other federal governments are supporting their research universities in supercomputer procurements, while in the United States, Federal support of university computer-based research is primarily directed toward procurement of discipline-dedicated minis and micros.

To see how the Foundation can play a vital role in helping the universities to assume their appropriate role, let us review several related NSF successes and accordingly the strengths they represent.

The development of computing within the universities was a direct consequence of the activities of the Foundation. The Foundation identified a need to develop university computing capabilities and programs, and backed that up with high priority funded programs in the late 1960's. This had significant impacts. The Foundation's approach of requiring matching funds resulted in a 3- or 4-to-1 multiplier of NSF grant dollars. The consequences of this leveraging of government funds accelerated the computer education of more students and generated more new computer-based research and applications ideas than any other Federal program. Our national computing leadership was obtained by direct and indirect interdependence between government, industry and universities. This achievement must certainly be judged in part as a consequence of these NSF funding policies. The three-legged stool approach was an operational success.

The National Center for Atmospheric Research is another success story. It was a bold step for the Foundation to establish a discipline-focused, government-funded and university-governed research organization incorporating a superscale computer facility. Today NCAR provides a respected cost-effective service to atmospheric scientists nationwide. We understand its service is now being expanded to the oceanographic community. To improve its service, it's developing a quality national network.

There are other unique strengths which NSF has which are well respected and have resulted, in our opinion, in the overall success of the Foundation's programs. The traditions of discipline-oriented funding, coupled with a strong peer review, are not duplicated in any other Federal program. This results in a broad-based research thrust with programs tailored to the university environment and national needs. Examples cover the full range of the sciences in pure research and now the industry-related research funded through the University Industry Cooperative Program.

When I read the Lax report and other documents which have been made available, I see the urgency for action. We, however, have yet to see an indication of plans for an NSF program for '84 and the years following. This is a concern because we are aware that it takes time to put a program in place and for the universities to respond. In addition to those universities who have installed supercomputers, the University of Minnesota, Colorado State University and Purdue University, we know that several others are keenly interested in planning supercomputer acquisitions. All are looking to the Foundation for leadership through operational support.

In identifying the attributes for a NSF supercomputer program, we have endeavored to match the needs of the university community and the strengths of the Foundation. The first attribute, therefore, is that the program should build on the Foundation's experience and the existing base of proven capabilities. We see some definite challenges here. The key issue is between a discipline-focused approach which you currently are so successfully applying and a more university-focused approach as used in the mid to late 1960's. A combination of the two appears to be necessary to meet the overall needs.

The second attribute is that the Foundation should identify supercomputer access for university-based, scientific research and education as a national priority.

Due to the pressing needs, the third attribute is that the Foundation should act quickly with the initial focus on providing access to existing sites via existing networks. Initially, access should be provided for key disciplines. We have identified a few based on our perspectives:

1. Computer Science - architecture/design of hardware/software, numeric/non-numeric algorithms, software automation, software quality assurance, artificial intelligence.
2. Engineering Sciences - VLSI design, CAD/CAM - electrical and mechanical, robotics, modeling and simulation, process control.
3. Physical Sciences - modeling and simulation.
4. Bio-Sciences - genetic engineering.
5. Economics - modeling.

The fourth attribute is long-term oriented. A plan should be developed and implemented for multiple university supercomputer sites utilizing a diversity of networked computer systems-vector, parallel and multiprocessor configurations.

This plan should include NSF funding of facility operating costs in part - drawing on previous successes in leveraging Federal dollars. Program support including access equipment for institutional researchers should be provided. The incorporation of a high performance networking capability can be achieved by drawing on existing resources and expertise.

The last attribute we see for this overall program is to foster cooperation between universities, government and industry. Perhaps the current University Industry Cooperative Research Program could be used as a model upon which to build inter-organizational ties to achieve common interests while leveraging investments.

We are reluctant to develop to the level of implementation the details of a program recommendation to support supercomputing research. However, we will speak of details to the extent that potential areas of cooperation can be identified or when awareness of interests may be of mutual benefit. Implementation considerations are of immediate and long term nature.

For the immediate, I have identified several program areas for research in applications. To become effective quickly in these areas, researchers will immediately require access to existing super-scale processors. We understand that the Department of Energy has recently made available 800 hours of supercomputer resource for energy-related researchers external to its sites. The Foundation should encourage similar cooperative commitments from other Federal agencies which will enable academic researchers access to supercomputer resources.

The Foundation can most quickly become a formal advocate by allocating funds from its resources to enable researchers access to one or more of the three installed university supercomputer sites. This firm commitment by the Foundation can in turn lead to further cooperation from industry including Control Data Corporation. We have already discussed the possibility with some NSF staff of supercomputer resource grants. We are prepared to speak more on this following our presentation.

Network access also requires immediate and special attention. Existing capabilities should be utilized whenever possible whether it be Federal, or private. Perhaps EDUNET or ARRANET can play a role. Interconnection of neighboring sites to share existing networks is also a consideration, for example, interconnection of the Colorado State University and NCAR supercomputer facilities.

The NSF program must include providing local hardware and software to enable researchers remote to the central supercomputer sites access to those sites. This idea presupposes one or more generally available access points on campuses which are remote to the supercomputer sites. To address this need, the Foundation's plans must incorporate a process for working through the remote university computing centers.

Guidelines on the local resource requirements will be necessary to minimize duplication of software development efforts.

Experience from the NCAR networking effort can be helpful in this area. To demonstrate Foundation commitment to the academic community, several pilot remote sites should be established in the upcoming fiscal year.

These are immediate considerations. For the longer term, the Foundation needs to establish a mechanism to plan and partially fund the required university supercomputer facilities and also to plan and implement a network access mechanism. I will discuss these in order.

For planning purposes, we contemplate a continuation of the general purpose super-scale vector machines. As noted earlier in the presentation, other architectures are also emerging. We consider the parallel processor, for example, as a special purpose very large or super-scale machine which has immediate application in artificial intelligence including image processing and in simulation. Control Data has had a machine of this type operational for five years. Lloyd Thorndyke will address the anticipated commercialization of that machine in his presentation. It is appropriate for Foundation planning to anticipate several machines of this type in universities.

The current distribution of machines among universities (or dedicated to their use) neglects critical regions of the country. This distribution exhibits the consequences of the lack of a national plan.

The Foundation could proceed by identifying regions for supercomputer systems and then urging universities in these regions to affiliate to promote cooperation in system funding and use. The UCAR model is a candidate for such affiliations. A mechanism to encourage an evolutionary growth of installed machines so as to maintain effective utilization is highly desirable. The Colorado State University model for building a user network utilizing a diversity of machines warrants serious consideration in that regard. An overall growth in university cooperation should be a strived for goal of the implementation.

As a candidate plan, the Foundation might consider the installation of three additional supercomputers in universities each year for two years. Planning and procurement for the first machines would begin in fiscal year 1984 with first installation at the end of that year or early in fiscal year 1985.

We all realize that universities are not in the relatively strong fiscal condition they were in the mid to late 1960's. Nevertheless, it is reasonable to expect a leveraging of Federal funds with state and private contributions to support university supercomputer installation and use. The magnitude of the matching funds could be initially determined by the number of sites required and the availability of other funds. Presumably, the computer resources purchased by the Foundation funds would be allocated, in a fashion similar to the NCAR model, to research programs for use. We are prepared to discuss funding requirements and models for host supercomputer sites.

Network planning and implementation is also complex. Perhaps existing networks as they are augmented can meet the needs. New emerging networks are also candidates. In the long term perhaps a dedicated network could be implemented in part with industry support and managed by a non-profit organization. The Foundation could provide support funds based on project use.

It's important to note here another NSF success not mentioned earlier. The Computer Science Network funded at a cost of $5 million may be an instructive model.

Certainly the great sophistication we all anticipate in network capabilities will open new opportunities for cooperation in research and networked supercomputers. We should continually strive to better leverage our resources utilizing this and other mechanisms available to us.

In summary, our message today is we believe the Foundation should champion a national program for research and development in super-scale computing applications.

I have identified and spoken about the following attributes of that program:

1. Build on experience and the existing base of proven capabilities.
2. Identify, as a national priority, supercomputer access and application research.
3. Start quickly with initial focus on access to existing sites/networks for key disciplines.
4. Develop and implement a long term plan for a national, multiple university supercomputer network.
5. Foster cooperation between government, university and industry through this process.

At the beginning of my talk, I summarized our message calling it a recommendation, a challenge, or, more appropriately, a call for commitment. In conclusion, let me speak to that last classification. I have noted that there is a general high level of awareness that something must be done to address the recommendations of the Lax report. That awareness permeates Federal agencies, universities and industry. Specifically, the academic community is waiting and watching - I know the environment well. Their expectation level is high that the Foundation will take the steps to establish a strong program in supercomputing. With academic expectations very high and their attention focused right here, the burden of responsibility is clear. We urge you to act quickly. The time is now.

CRAY RESEARCH, INC.

STATEMENT OF

MR. JOHN A. ROLLWAGEN
CHAIRMAN AND CHIEF EXECUTIVE OFFICER
CRAY RESEARCH, INC.

Before the Subcommittee on Energy Development and Applications
and the Subcommittee on Energy Research and Production

June 14, 1983

MR. CHAIRMAN AND MEMBERS OF THIS DISTINGUISHED COMMITTEE. I am John A. Rollwagen, President, Chairman of the Board and Chief Executive Officer of Cray Research, Inc., the world's leading supplier and developer of supercomputers and supercomputing technology. Since it was founded in 1972, Cray Research has grown to over 1400 employees located across the United States and in five foreign countries (United Kingdom, West Germany, France, Netherlands, and Japan). By the end of 1983, we will have added two more countries -- Canada and Sweden -- to our foreign installed base. We currently manufacture four major products, including two types of central processors, the CRAY-1M and the CRAY X-MP. The CRAY-2 is in final stages of development and will be available in late 1984. To date, we have installed fifty-two CRAY-1 supercomputers, approximately eighty percent of all current generation supercomputers worldwide. By year end, we expect to install 14 new systems including six of the new CRAY X-MP's.

Thank you for this opportunity to speak with you concerning the current and future status of the supercomputer development required to meet future challenges in energy research. We appreciate and commend the concern you have demonstrated by these hearings. We also appreciate your broader concern that the United States maintain its strong position of world leadership in supercomputing technology. While I am happy to report that we at Cray Research are excited and optimistic about the future, there are some matters of deep concern which could place us at a considerable disadvantage if they are not dealt with expeditiously. I shall discuss these concerns later in this testimony.

For the purpose of this testimony, "supercomputer" refers to very high speed general purpose computer systems such as those produced by Cray Research and used by the Department of Energy Laboratories at Los Alamos and Albuquerque, New Mexico; and at Livermore, California. We are specifically not referring to Artificial Intelligence or "Fifth Generation" systems which are the subjects of a major new research program recently proposed by the Department of Defense Advanced Research Project Agency (DARPA). Since there has been considerable confusion in the press, this distinction is important.

IMPORTANCE OF SUPERCOMPUTERS

BASIC TOOL FOR ENERGY RESEARCH

Almost forty percent of our systems are used to support some type of energy research. Most of these systems support energy research conducted at DOE National Laboratories, and similar facilities in Germany, France and the United Kingdom. Other witnesses at these hearings have already detailed the important contribution that supercomputers are making at these laboratories -- both by the solution of otherwise intractable problems and by allowing the laboratories to conduct their research at a dramatically reduced cost by eliminating the need for many physical experiments.

A second area of energy research to which supercomputers are the key is that in support of exploration for and development of petroleum resources. Beginning with Atlantic Richfield Oil Company, which acquired a CRAY-1 in late 1980, seven major oil companies have installed supercomputers. Using advanced modeling and seismic signal processing techniques, oil companies expect to demonstrate dramatic dollar savings in exploration and reservoir management. These savings translate directly to the American consumer; better reservoir management allows us to conserve our petroleum reserves.

Finally, in addition to the direct energy research, there are a number of related opportunities which are highly dependent upon current and future generations of supercomputers. One of the most important lies in the area of Computational Fluid Dynamics, discussed earlier by Dr. Kerrebrock, of NASA. In addition to the significant impact upon the entire process of engineering design, recent results from the use of the CRAY-1 supercomputer by the aerospace community have demonstrated the potential for designing air frames with significant savings in total fuel consumption, translating into direct savings to both the public and private sector and preservation of national fuel reserves. We expect the impact of these savings to grow as supercomputers are introduced into the automobile and shipbuilding industries.

FUNDAMENTAL TOOL FOR DEVELOPMENT OF NEW TECHNOLOGIES

Supercomputers have made significant and well-documented contributions to energy and general scientific research. However, we have only begun to realize the potential impact that advanced computing could have upon our Nation's economy and security. The results we have obtained using this generation of computers have spawned new methods fundamental to the birthing of emerging technologies. Some of these technologies have already reached the National agenda, such as the recently announced Presidential initiative to design a space-based anti-ballistic missile system. Access to state-of-the-art systems will be key to advances in technologies of societal and political importance including biotechnology and medical research. Recognizing these facts, other Western countries are moving to capitalize upon the potential that supercomputing offers. The French have instituted a national supercomputer development project and begun to broadly apply supercomputers. The British are not far behind, both in developing their understanding of supercomputing and in preparing for national initiatives. This committee is well aware of Japanese efforts.

CONTINUED WORLD LEADERSHIP A PRIORITY

Mr. Chairman, the world is on the verge of a major computing revolution as supercomputing is applied across a broad industrial and scientific base. The only question at this juncture is which nation will choose to lead this revolution.

The cost trade-offs are dramatic. If an oil company can be spared drilling a single dry hole, the entire supercomputer installation will be paid for several times over. If a major aircraft or a new automobile model can be made as little as 2 or 3 percent more fuel-efficient, the savings to the nation in domestic fuel reserves can be staggering. We have not yet begun to reap the benefits the potential of the use of supercomputers in biotechnology, medical health care delivery systems, socio-economic modeling and a vast array of operational defense problems. As we push more into space, the importance of processing huge amounts of data in a short time become even more important. Supercomputers have already been shown in some evaluations to have considerable efficiencies and capabilities not realizable in any other class of system. Even in the area of artificial intelligence, which is a subject of considerable research around the world, supercomputers stand to be a powerful tool. Artificial intelligence algorithms are known to be very complex, but the payoff in certain areas could be substantial if these algorithms can be developed in a reasonable amount of time.

As we approach the environment of the year 2000, we should not be surprised that Japan, France and the United Kingdom have all recognized the importance of supercomputing to the extent that they have initiated national programs of varying scopes to deal with the issue. Many have questioned why the Japanese would enter supercomputing since their normal strategy is to target industries with broad commercial significance. While it is possible that the venture into supercomputing has been more for prestige, I believe that history will record the venture as a very shrewd one. When Fujitsu and Hitachi announced their supercomputers last year they projected deliveries to approximately 30 domestic customers in the first 18 months, or more new customers than Cray Research expects worldwide during the same period. At this time there are only two current generation supercomputers in Japan. If these Japanese manufacturers are successful, we can expect to see a computing revolution throughout Japanese industry that could further threaten U.S. industrial competitiveness.

As these facts indicate, the importance of leadership in supercomputing technology goes far beyond the already well-established links to Defense and National Security. As I said earlier, we are on the verge of a major worldwide computing revolution in which supercomputers stand to play a fundamental role. Leadership in supercomputing technologies, and in the aggressive implementation of that technology, could play a significant role in the future of both our basic industries, and new high technology frontiers.

Cray Research is the world leader in supercomputer technology, not merely from a desire to be better than anyone else, but because we are seriously addressing important problems, such as those found in energy whose solutions are of National consequence. We are addressing these problems in a constructive, imaginative and American way -- and the result has been excellence. The United States as a Nation and Cray Research as a private enterprise have each demonstrated their ability to achieve world leadership. But leadership does not guarantee success. True success is sustained achievement. Our challenge today is to ensure that we can continue to keep that leadership. In meeting that challenge, we must not focus upon the plans or accomplishments of others. Our primary motivation is not to defeat the Japanese. For the Nation and for Cray the current challenge is to meet society's needs. In order to adequately address the future needs of our industries, our technology, and the National Security, we must foster an environment of dynamism in research and development. We must define the problems to be solved and focus upon the solutions rather than the competitors.

Thus far, we are confident that the United States is still in the leadership position in supercomputing at this time. Although some of the announcements by Japanese manufacturers have indicated impressive capabilities on paper, the real performance test for supercomputers is their capability to get the scientific job done. It will still be quite a while before that can even be evaluated, but we expect that the current and near-term future offerings of Cray Research will equal or surpass the Japanese challenge.

Cray Research, and, we believe the Nation, is ready to meet the supercomputing needs of our research and industrial community. In order to best accomplish our work however, we need an atmosphere which will continue to spark and reward the enterprise which has fostered the growth of our industry for the past twenty-five years. The formula for success is far from simple. Until recently, the success of Cray Research has depended largely upon factors within our control, such as management style, internal commitment to research, etc. We have worked hard to optimize those terms of the formula. We have promoted innovative management. We have committed 15-17% of our gross annual revenues to research and development, yielding a total supercomputer development budget comparable to the entire published budget for the Japanese Supercomputer Development Project. We have created an environment where a other supercomputer designers can flourish side by side with Seymour Cray. We have marketed aggressively to foreign customers located within Allied nations to the point that over forty percent of our new revenues are coming from foreign sales. We have sought to establish and maintain close contact with our customer base to ensure that we are sensitive to their future technology needs. These efforts have proven fruitful thus far, and we believe we have secured supercomputing preeminence for the next generation of computers.

Today, however, other terms of the formula -- national economic and trade policy, semiconductor availability and domestic supercomputer policy -- have risen in importance. In addition, the specter of successful industry targeting by the Japanese, French, and British has dramatically increased the potential negative impact of any artificial barriers imposed by Federal policy.

NEED FOR TECHNOLOGY-ORIENTED ECONOMIC AND TRADE POLICY

With the serious challenges from the Japanese, British, and French, many of our efforts hinge upon technology-related economic policy. Since we devote large financial resources to research, our efforts have benefited substantially from recently enacted temporary research and development tax incentives. These incentives need to be strengthened and enacted permanently. Because our financial success and our position of world leadership is dependent upon foreign sales, we are very sensitive to changes in export administration and West-West technology transfer policy. We need a consistent, well-managed, and reasonable policy which will protect National security without harming our National economic assets. Finally, efforts to enhance the quality and quantity of trained scientists and engineers will all have a substantial impact upon the future of U.S. leadership in supercomputing. These issues have been the subject of other hearings, and I shall not address them further today.

NEED FOR ADVANCED SEMICONDUCTORS

Lack of high speed semiconductors is currently the Achilles heel of the U.S. supercomputer effort. Product development projects have been delayed or even suspended due to the lack of new high-speed reliable semiconductor parts. This is particularly true concerning logic chips which we are very reticent to buy from the Japanese. We already purchase most of our memory chips from abroad. We are concerned that we will have to rely upon the Japanese semiconductor industry for the special chips required to build our computers. If this happens, U.S. leadership would begin to depend upon the willingness of Japanese suppliers to sell us the most advanced chips at the same time that these chips are available to our Japanese competitors. Our second concern with this situation is that if we buy our logic chips from the Japanese we will be forced to transfer to our supplier the proprietary mask design which reveals exactly how we will be using the chip. This type of technology transfer could be detrimental to both U.S. competitiveness and National Security.

Finally, we are concerned that the major government sponsored research is not addressing the semiconductors most useful to high speed semiconductors. We are not hopeful that either VHSIC or the DARPA Galium Arsenide programs will yield any advantage to the U.S. supercomputer industry.

Having addressed what we consider the primary impediments to U.S. leadership in supercomputing, I will now turn to the final area of the "formula for success" -- that of the Federal role in fostering supercomputer development.

THE FEDERAL ROLE IN FOSTERING SUPERCOMPUTER DEVELOPMENT

CURRENT STATUS OF SUPERCOMPUTER DEVELOPMENT

We are currently in an exciting phase at Cray Research. By the end of next year, we will have delivered a major new product each year for four years. In 1981, we delivered the first Input/Output subsystem, a special high speed processor which dramatically enhanced the capability of the CRAY-1 supercomputer. In 1982 we delivered the first CRAY X-MP parallel processing supercomputer to our internal data center, marking the first time in over twenty years that someone other than Seymour Cray had designed a commercially successful supercomputer faster than Cray's latest system. We also delivered the first Solid State Storage Device, a high speed, 256 million byte, auxiliary memory which connects to the

CRAY X-MP via a communications link capable of operating at speeds up to ten gigabits per second. In 1983, we will deliver the first CRAY 1/M, a follow-on to the CRAY-1 processor which uses MOS (instead of ECL) memory, thus reducing the system price as much as fifty percent. In 1984, we expect to deliver the first CRAY-2 system, which will usher in a number of new technologies required to meet our goals for the second half of the decade.

We are very encouraged with the current state of supercomputer development. If the market continues to be robust, and if some of the software and applications difficulties can be overcome, it is not overly optimistic to assume that the system which Cray Research delivers in 1990 will perform at speeds at least two orders of magnitude faster than the CRAY-1; memory sizes are expected to increase at the same rate.

In the applications area, we are not quite as encouraged. The record of the last five years indicates that other than in environments such as the Department of Energy Laboratories, the U.S. research and engineering community has been slow to implement supercomputing technology. Even the Defense and aerospace communities, which have in the past aggressively applied new computers, have lagged seriously behind their counterparts overseas. Since the success of new computing technologies is heavily dependent upon their application in a user environment, lack of aggressive use of new generations of supercomputers could impede attempts to speed up the development cycle.

In the software area, there is a need for parallel processing breakthroughs in software technology. There is a critical need to move the state of the art in this area as quickly as possible. Now that the CRAY X-MP is available, it is possible for researchers to simultaneously investigate parallel processing and high speed vector processing. Unfortunately, unless action is taken in the near future, this unique tool, like supercomputers before it, will be largely inaccessible to the vast University research community.

The greatest immediate challenge from the Japanese as well as from Western Europe is their aggressive of new technology. In the United States, supercomputers are largely inaccessible to the vast majority of defense, industrial and academic scientists and engineers. The picture is quite different in Europe and Japan, where the governments have long recognized the importance of supercomputers to their world competitive positions.

THE PROPER ROLE OF THE DEPARTMENT OF ENERGY IN SUPERCOMPUTER DEVELOPMENT

The Department of Energy Laboratories have played a crucial partnership role in the development of supercomputers since Mr. Cray began his career. In addition to providing an immediate market for successful products, the Department has sponsored software development and application research in an atmosphere of close cooperation with Cray Research. In many respects, this particular vendor-customer relationship has been a crucial key to the success of Cray Research's efforts in developing another generation of supercomputer beyond the original CRAY-1. I am aware that this Committee has encouraged the Department to continue in this crucial role and commend you for this action.

In his earlier testimony, Dr. Trivelpiece of the Department of Energy pointed out the importance of strengthening this partnership if we are to meet the challenges. We at Cray Research strongly echo this statement. We are entering a new generation of development in a number of dimensions. Ten years ago, when Mr. Cray was building the first CRAY-1, he was able to use supporting technologies which were immediately available. There was no foreign competition, and there was no indication that future development would be limited by availability of domestic parts. Furthermore, from a National Interest standpoint, four major U.S manufacturers were expending major resources to compete

(Burroughs, IBM, CDC, and Texas Instruments). Today, the picture is quite different. Lack of parts is impeding progress and the bulk of both spending and competition is coming from abroad. In addition, the importance of supercomputing to the next generation of industrial technology has elevated supercomputing from a specialized domain of energy research and a limited set of other customers to a technology fundamental to the competitiveness of U.S. industry.

Finally, although we applaud the efforts of the Department and others to assist in supercomputer development, we are concerned that this zeal could result in efforts to attack the wrong problems. Specifically, I offer the following recommendations as you consider the proper course:

° Ensure that the Department is sufficiently funded to pursue its charter. Using supercomputers, the Department is able to leverage the taxpayer investment better than with any competing technology. Under no circumstance should computer equipment budgets be curtailed in favor of development projects, or other research.

° Enhance the Department's capability to acquire computers strictly for software and algorithm research. The most advanced system available should always be available for general research. This should not be limited, of course, to "experimental" systems, but should include systems such as the CRAY X-MP or CRAY-2.

° Encourage and provide extra appropriations for University access under Department sponsorship. One way to facilitate this access would be to move CRAY-1 systems from the larger laboratories to smaller laboratories as the CRAY-1's are replaced by CRAY X-MP's, CRAY-2's or other systems.

° Encourage continued close partnership between Universities, industry and the laboratories. We are committed to work closely with the Department to define the best mechanism for this purpose.

CONCLUSION

In conclusion, Mr. Chairman, I am pleased to appear before you this afternoon as a representative of our company and our customers. Even more important, however, I am appearing as the representative of our technology and those who are most directly responsible for American dominance of this vital technology: Seymour Cray, the chief designer of the CRAY-1 and its predecessors, the Control Data 1604, 6600, and 7600; Les Davis, Cray Research Executive Vice President for Manufacturing, Research and Development, and the individual who has delivered the Cray machines to the users; and Dr. Steve Chen, the chief designer for the first parallel processing supercomputer, the CRAY X-MP, which is currently the world's fastest and largest capacity supercomputer. Individuals like these -- very select individuals -- develop supercomputers capable of commercial and scientific success. Development of successful supercomputers is an art and a discipline which appears to emanate from a "critical mass" of talent, single-mindedness, support, and the opportunity to win.

Although we may address issues of importance to individual companies or customer constituencies, the success of these hearings and any action which the Congress or Administration takes will ultimately be measured by how that action fosters an atmosphere in which these individuals, and others waiting to emerge can flourish and deliver new generations of systems to the open marketplace. If Cray Research can provide an environment where new creative talent can emerge, and at the same time provide an environment where an exceptional individual like Seymour Cray can continue to be creative, we will indeed have done something very special. We believe we have begun to reap the benefits of that accomplishment and we are committed to helping our Nation to do the same.